Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

A COMPARISON OF HAND TRACING AND COMPUTERISED CEPHALOMETRIC ANALYSIS

Caroline Mary Marron

SUBMITTED IN PARTIAL FULFILMENT OF THE CLINICAL DOCTORATE IN DENTAL SURGERY (ORTHODONTICS)

TRINITY COLLEGE DUBLIN

2021

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

____

Caroline Mary Marron

Summary

The aim of this study was to determine if automated computerised cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software is as accurate as traditional hand tracing in terms of landmark identification, angular and linear measurements.

Pre-treatment lateral cephalograms for a selection of patients undergoing comprehensive orthodontic treatment were collected and irreversibly anonymised. Using a sample of eight radiographs, an inter-examiner reliability investigation was carried out to measure the level of inter-rater agreement of cephalometric analysis by hand tracing between the primary investigator and two Consultant Orthodontists. Using a sample of thirteen radiographs, an intra-examiner reliability investigation was carried out to measure the level of intra-rater agreement for cephalometric analysis by hand tracing of the primary investigator against themselves at two different time points. To compare the results of automated computerised cephalometric analysis using Planmeca Romexis® (Version 5.1.1.R) dental imaging software with those obtained by traditional hand tracing, 68 lateral cephalograms were manually traced by the primary investigator to include 16 anatomical landmarks and 15 angular and linear measurements. The radiographs were then analysed by the automated software programme. The accuracy of the computer software programme was assessed by comparing the automated landmark locations and measurements to the hand traced landmark locations and measurements. Accuracy for automatic landmark identification was recorded as successful detection within a 2.0mm precision range of the hand traced landmark. Accuracy of automatic angular and linear measurements was recorded within 2° or 2mm of the hand traced measurements respectively. Intraclass correlation coefficients were calculated to measure the level of inter-rater agreement. Bland-Altman plots were constructed to analyse the level of inter-rater agreement.

Planmeca Romexis® (Version 5.1.1.R) dental imaging software programme had a landmark successful detection rate (SDR) in a 2.0mm precision range of 82.08% with a mean radial error of 1.25 ± 0.91mm. This is comparable with the inter-rater reliability between two clinical experts. Gonion and A-point were the least accurately identified landmarks. The average success rate for all angular and linear measurements within 2° or 2mm respectively was 80.98%. ICCs for all measurements ranged between 0.73 and 1.00 indicating moderate to excellent levels of agreement.

Automated cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software is suitable for clinical use, however, more accurate measurements may be possible if landmark locations are checked and amended if necessary prior to performing an analysis.

Acknowledgements

I would like to sincerely thank my research supervisors Dr Angus Burns and Dr Thérèse Garvey for all their help, support, encouragement, and patience throughout this project.

I would also like to thank Dr Timothy McSwiney for his contribution to and support of this project.

Thank you to Ms. Isabel Fleischmann, librarian in Dublin Dental University Hospital, for her assistance with the literature search strategy.

Thank you to Dr Erica Donnelly-Swift and Dr Maria Van Harten for their statistical expertise and guidance for the data analysis.

Many thanks to the patients and parents who consented to participate in this study, this project would not have been possible without them.

I would like to thank Emily Higgins for her friendship, support, and encouragement over the last three years.

Finally, I would like to thank my friends and family, particularly my husband Eric, my sister Louise and brother-in-law Dave and my parents, Margaret and Alfie, for their unwavering support over many years and throughout this programme.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AAM	Active Appearance Model
AI	Artificial Intelligence
ANS	Anterior Nasal Spine
ASM	Active Shape Model
CNN	Cellular Neural Network
CVM	Cervical Vertebral Maturation
CBCT	Cone Beam Computed Tomography
FALA	Fully Automatic Landmark Annotation
FMPA	Frankfort Mandibular Plane Angle
ICC	Intraclass Correlation Coefficient
IEEE	Institute of Electrical and Electronics Engineers
ISBI	International Symposium on Biomedical Imaging
MMPA	Maxillo Mandibular Plane Angle
MRE	Mean Radial Error
PNS	Posterior Nasal Spine
SDR	Successful Detection Rate
YOLOv3	You-Only-Look-Once version 3

Chapter 1 Literature Review

1.1 Literature Search Strategy

Literature examining automated cephalometric analysis and automated identification of anatomical landmarks on two-dimensional (2D) lateral cephalogram radiographs was reviewed. A literature search was carried out using PubMed, ScienceDirect and Google Scholar. Reference list searching of included studies was performed and grey literature publications were also included. The search strategy is included in Appendix 1 and a flow diagram representing the literature search is shown in Figure 1.1.

1.1.1 Inclusion Criteria

Studies related to:

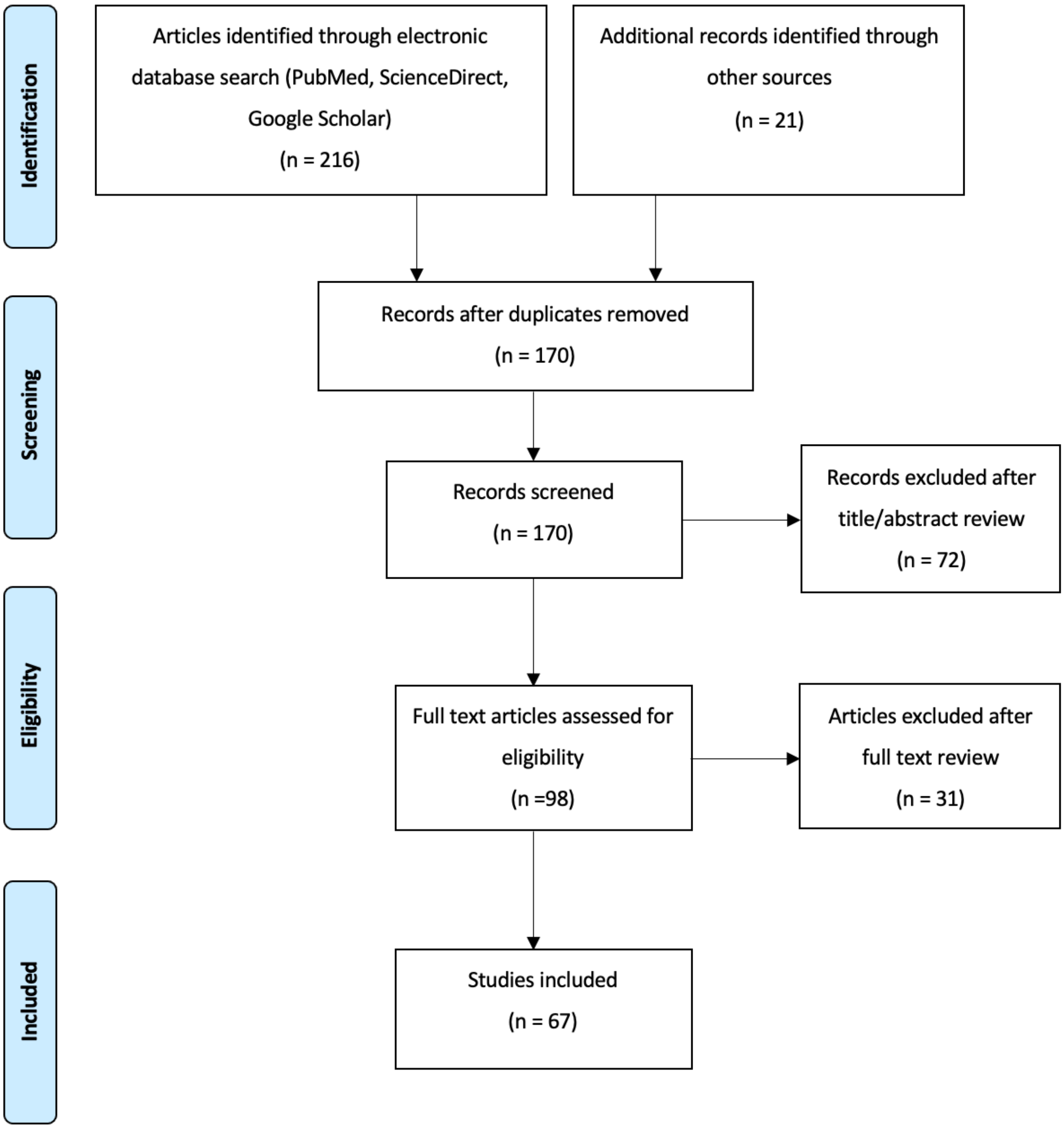
- Automatic identification of anatomical landmarks on 2D lateral cephalograms
- Automated cephalometric analysis of 2D lateral cephalogram

1.1.2 Exclusion Criteria

Studies related to:

- 3D imaging, most commonly cone beam computed tomography (CBCT)
- Automated analysis of other types of radiographs e.g., panoramic radiographs and posteroanterior (PA) cephalometric radiographs
- Automated assessment of facial aesthetics or assessment of facial asymmetry
- Automated measurement of craniofacial deformities
- Automated Cervical Vertebral Maturation (CVM) assessments
- Automated analysis of post-surgical facial changes
- Automated cephalometric superimpositions
- 3D printing

Figure 1.1 Literature Search Strategy[1]



1.2 Cephalometric Analysis

The cephalometer was introduced by Broadbent and Hofrath in 1931[2] and since then cephalometric analysis has become a standardised diagnostic method in orthodontic practice.[3, 4] Traditional cephalometric analysis is carried out by hand tracing a plain film radiograph, which involves anatomical landmark identification followed by angular and linear measurements.[5] These measurements are used primarily as an aid to orthodontic diagnosis and treatment planning. While traditional cephalometric analysis is a valuable diagnostic tool, this method is time consuming and can result in errors due to clinicians' subjectivity.[6, 7] Since the 1980s, there have been many efforts towards simplification of cephalometric analysis with the help of computer-based technology.[8]

1.3 Digital Radiography

The introduction of digital radiography has several advantages including the ability to obtain an instant image, reduced radiation doses for patients and the facility to enhance the image quality.[9-11] Digital imaging in relation to lateral cephalograms has introduced other potential advantages including the ability to carry out automated or semi-automated computerised cephalometric analysis, the capability to digitally superimpose lateral cephalograms and the facility to use digital cephalometry as an aid to the planning of surgical cases.[12] The widespread use of digital radiography led to the development of computer-aided cephalometric analysis. Early cephalometric software programs had the capability for automatic angular and linear measurements following manual identification of anatomical landmarks. More recent software programs allow for automated cephalometric analysis, which involves computerised landmark identification and measurements. This was introduced with the aim of reducing the time required to obtain an analysis, improving the accuracy of landmark identification and reducing operator errors.[13]

1.4 Hand Tracing

The reliability of hand tracing has been studied extensively and, while it is associated with random errors, particularly in landmark identification [7, 14], it is still considered the gold standard when carried out by an experienced clinician under the right conditions.[8, 15] The ideal conditions for performing a cephalometric analysis are that it is done in a dark room with a light box, using an acetate x-ray, on a sheet of tracing paper with adhesive tape to secure it and using a 3H pencil. The radiograph should be oriented with the Frankfort plane horizontal, and the tracing paper should be taped to the acetate x-ray in a way that allows it to be lifted during the tracing to check obscure points without disrupting the position of the tracing paper. During the tracing process, specific anatomical landmarks are identified, and these landmarks form the basis of angular and linear measurements that are recorded and used to assess the dental and skeletal relationships.

1.5 Accuracy/Reliability

Errors in cephalometric analysis are described in two categories; systematic errors or random errors .[5] Systematic errors occur when there is variation in the size and scale of the radiographic images and this is not taken into consideration and compensated for during the analysis.[16] Random errors can occur during any stage of cephalometric analysis, from tracing and landmark identification to measurement errors when calculating angular and linear measurements.[5] The greatest and most significant source of random error is in the identification of anatomical landmarks.[16] This source of random error is common to both manual and computerised cephalometric analysis.[13] Previous studies have found that the accuracy of landmark identification varies considerably depending on the operator, the landmark and the quality of the image.[17-19] Furthermore, as landmarks are currently defined as descriptions rather than mathematical conditions, localization is subjective.[19]

The main challenge of accurate landmark identification relates to the superimposition of anatomical structures and the projection of a complex 3D structure onto a 2D film.[20] The accuracy of individual landmark identification is related to the shape of anatomy being examined. Where the landmark lies on a well-defined junction between two edges, it is more easily identified, e.g., the upper incisor edge.[3] However, where the landmark lies on a curved edge there is more ambiguity as to its precise location, e.g., A-point, B-point, pogonion and menton.[3] The landmarks most associated with errors when manually identified are gonion and lower incisor apex and the most accurately located landmarks are sella and nasion.[3, 14] Gonion is a constructed point and is therefore more difficult to reproduce accurately than other landmarks. Gonion can be affected by asymmetric overlapping of the right and left sides of the mandible, which is exaggerated where there are head positioning errors in the cephalostat or in patients with mandibular asymmetries.[3, 7] It is not uncommon to have two lower borders of the mandible visible on the lateral ceph, which can affect the ability of the operator to accurately construct this point. The lower incisor apex is also frequently associated with errors as the surrounding "noise" in that area of overlapping structures is so great that often the position of the lower incisor apex has to be estimated based on the angulation of the crown and the likely shape and length of the root.[3]

In terms of margins of error, Forsyth and Davis report that an error in landmark localisation of no more than ±1mm is desirable.[12] Rakosi in 'An Atlas and Manual of Cephalometric Radiography' published in 1982, stated that an error of ±2mm is acceptable in relation to landmark localisation,[21] which is a view that is widely agreed upon and reported in the literature. The mean error of expert landmarking has been reported as 1.26mm.[3, 18]

The acceptable measurement error for angular measurements is no more than 2° and for linear measurements is no more than 2mm.[8, 22, 23] Measurement error is affected by human error, the limitations of the measurement tool e.g. rulers, protractors and also by the limitations of human vision, which varies between operators[24-27].

1.6 Are lateral cephalograms necessary?

There is some debate in the orthodontic community as to whether lateral cephalograms are indicated or not. Cephalometric analysis was first developed to assess craniofacial growth,[2] not for orthodontic diagnosis, and so it is an imperfect diagnostic tool that is associated with both systematic and random errors.[3] Head positioning errors, magnification errors, operator errors and landmark identification errors can cause problems with the accuracy of the analysis.[7]

Several studies have been undertaken to assess the need for lateral cephalograms in orthodontic diagnosis and many of these have concluded that cephalometrics are not required as clinical examination and study models are sufficient for treatment planning.[28-31]

However, despite these conclusions, cephalometric analysis is still a widely used orthodontic diagnostic tool. The British Orthodontic Society Guidelines for the Use of Radiographs in Clinical Orthodontics provide guidance on when lateral cephalograms are and are not indicated.[32] A lateral cephalogram is indicated to assess the skeletal pattern and labial segment inclination pre-orthodontic treatment and also to monitor the effects of treatment.[32] The guidelines state that if the incisor relationship does not require significant change or if only an upper fixed appliance is being prescribed that lateral cephalograms are unlikely to be required but that they are indicated for patients who have a significant skeletal discrepancy.[32] Lateral cephalograms, and any other radiographs, should only be taken when clinically justified to avoid unnecessary irradiation of patients.[32]

1.7 Semi-automated Cephalometric Analysis

Semi-automated cephalometric analysis involves manual identification of anatomical landmarks followed by automatic angular and linear measurements. The accuracy of semi-automated cephalometric analysis was examined in the 2000s and several studies were carried out comparing traditional hand-tracing to computer-aided analysis. Computer-aided analysis has also been associated with errors in landmark identification[33] but, overall, studies have found that the two methods are comparable in terms of clinical reliability.[23, 34] However, traditional hand tracing performed by an experienced clinician under the right conditions is still considered the gold standard as variability between various semi-automatic software programmes can influence accuracy of landmark identification,[8, 15] which could be a confounding factor in cephalometric research.

1.8 Fully Automated Cephalometric Analysis

The first attempt at automated computerised cephalometric analysis was undertaken by Cohen et al in 1984.[35] They assessed the accuracy of locating two landmarks, sella and menton, on 23 radiographs using an image processing system called CLIP4. They had limited success with this method; however, it demonstrated the possibility of automated cephalometric analysis, and it has been succeeded by the development of many different techniques attempting to automate this process over the past 37 years with steadily increasing success. There is limited research currently available on these software programs and further research is needed to assess their accuracy and reliability.

1.9 Measurement Methods

As described in the literature, to measure the accuracy of fully automated cephalometric analysis the computer software program is compared to a human expert, either by traditional hand tracing and measurements or by manual digital identification of landmarks and digital measurements. Some studies report that all the landmarks and measurements were recorded by one clinical expert[36-39] and other studies report calculating the mean landmark location of multiple clinical experts following landmark identification of each expert on all of the study radiographs.[12, 40-43] The accuracy is measured as the difference between the clinical expert measurements, i.e., the gold standard for the study, and the automated measurements.

In the available literature to date, the main primary outcome measures recorded were successful detection rate (SDR) in a 2mm precision range, the mean radial error (MRE) and the classification accuracy as defined by the description of each anatomical landmark.[44]

Several studies also reported the x- and y-axis coordinates of the landmarks to enable assessment of the direction of error if applicable.

1.10 Technical Approaches Used to Automatically Identify Cephalometric Landmarks

The various techniques for computer software programs capable of detecting anatomical landmarks have been broadly divided into five main categories. [5, 45, 46] These techniques are largely in chronological order in relation to sequence of development and testing. The five categories are:

1. Image filtering plus knowledge-based approaches
2. Model-based approaches
3. Soft-computing or learning approaches
4. Hybrid approaches
5. Artificial Intelligence (AI) based models

The literature related to each of these categories will be discussed in detail in this review.

1.11 Image Filtering and Knowledge-Based Approaches

Early techniques were based around image filtering and knowledge-based approaches. Table 1.1 summarises studies that used these approaches to automatically detect anatomical landmarks on lateral cephalograms and were conducted between the mid-1980s and the late 1990s. This method was first described by Lévy-Mandel in 1986[47], and it laid the foundation on which other researchers built to refine this technique. It was based on the idea that during hand tracing, specific anatomical features of interest on a lateral cephalogram are visualised and traced as either lines or edges prior to identification of individual landmarks[47]. For a computer software program to identify these lines and edges, they were highlighted first by image filtering and then by edge enhancement to track the edges of well-defined anatomical borders, e.g., the lower border of the mandible. This process is known as line enhancement. The relevant lines and edges were then identified using a knowledge-based line-tracking algorithm. This knowledge was then used to identify nearby landmarks. Lévy-Mandel et al. (1986) successfully detected 23 out of 36 landmarks on two good quality radiographs using a knowledge-based line algorithm.[47] The accuracy and mean error were not reported in this study. Knowledge-based edge detection approaches work well for landmarks that are located on or in close proximity to well-defined edges. However, these methods do not work well in the context of less well-defined anatomical landmarks.

A resolution pyramid is a series of copies of an original image in which the resolution of each copy is sequentially reduced[18]. The highest level of the pyramid, or in other words the lowest resolution image, is assessed first to identify the main relevant features and then the higher resolution images are assessed, if required, to carry out a more refined search[18, 48]. The benefit of this is a substantial reduction in computation time.[48] Parthasarathy et al. (1989) used resolution pyramids and a knowledge-based hand-crafted algorithm to locate 9 landmarks on 5 lateral cephalograms and demonstrated successful detection of 58% of the landmarks within 2mm with a mean error of 2.06mm.[18] The limitations of this study were related to the small sample size and the rigidity of the rules associated with knowledge-based algorithms making it very difficult to add new landmarks and also being unable to locate soft tissue or complex structure landmarks. This work was followed by Tong et al (1989) who also used resolution pyramids and an improved knowledge-based hand-crafted algorithm to locate 17 landmarks on five lateral cephalograms.[48] They reported successful detection of 76% of the landmarks within 2mm and a mean error of

1.33mm. This method was more computationally efficient than previous methods but reported an issue with accuracy on the posterior mandible and the ascending ramus where only 44% of landmarks were located correctly.[48]

Knowledge-based approaches are dependent upon strict rules and therefore rely on the quality of the radiographs. The accuracy of landmark detection significantly reduces if the image quality is compromised, for example, if radiographic artefacts or structural anomalies are present.[12, 18, 47, 49] Therefore, subsequent studies refined the technique to try to improve the image quality as much as possible,[12, 18, 48, 50] however the limitation in relation to image quality could not be overcome. This approach to automatic landmark detection is inherently flawed in this respect and is not fit for purpose in the context of the clinical application of these software programs. Furthermore, in two of these studies, testing was carried out on the training set, which is an experimental design flaw.[18, 47]
Later studies by Forsyth and Davis (1996) and Ren et al. (1998) used knowledge-based approaches and reported increased accuracy of these algorithms,[12, 50] as shown in Table 1.1, however they were found to not be as accurate as manual landmark identification and could not be recommended for clinical use.

Overall, image filtering and knowledge-based approaches were unable to compete with traditional manual identification of anatomical landmarks in terms of accuracy and reliability.

Table 1.1 Summary of the accuracy reported by studies using image-filtering and knowledge-based approaches

Author (Year)	No. of test X-rays	No. of landmarks and accuracy	Techniques	Comment
Lévy-Mandel (1986)[47]	2	23 out of 36 landmarks identified successfully Accuracy and mean error were not reported	Edge enhancement Knowledge-based line extraction	Only successful with good quality images Criteria for successful location of landmarks were not specified Comparison method not described Very small sample size
Parthasarathy (1989)[18]	5	9 landmarks 18% < 1mm 58% < 2mm 100% < 5mm Mean error = 2.06mm	Resolution pyramid Knowledge-based line extraction	Analysis took 10 minutes to perform – speed greatly reduced compared to Levy-Mandal. Does not work if artefacts present or if poor quality image – cannot be used clinically Small sample size
Tong (1989)[48]	5	17 landmarks 76% < 2mm Mean error = 1.33mm	Resolution pyramid Edge enhancement Knowledge-based extraction	Extended the work of Parthasarathy (1989)[18] Criteria for successful location of landmarks were not specified Small sample size
Forsyth and Davis (1996)[12]	10	19 landmarks 63% < 1mm 74% < 2mm	Gray level value difference Edge location Knowledge-based extraction	Excluded x-rays where the apices of the incisors were obscured by unerupted teeth
Ren (1998)[50]	10	24 landmarks < 1mm 19 landmarks < 0.5mm 5 landmarks not identified	Image enhancement using image layer	Comparison method not described

1.12 Model-Based Approaches

Table 1.2 summarises studies that examined the ability of model-based approaches to automatically locate anatomical landmarks. These studies were conducted between the mid-1990s and 2010. Model-based approaches were first reported in the context of automated cephalometric analysis by Cardillo and Sid-Ahmed in 1994.[51] They proposed a new concept in this field, which aimed to locate the landmarks directly rather than replicating the orthodontist's method of tracing the characteristic anatomical features and then subsequently identifying the landmarks. This was carried out using pattern recognition, also known as target recognition. A model of the important anatomical structures is used to reduce the search area for each landmark and then shape-matching techniques are used to precisely locate the landmarks of interest. This model is then used as a template to locate landmarks on other images. It is adjustable and can be matched to equivalent areas to allow for landmark identification.[36] The cephalometric target recognition algorithm can be trained to recognise the shape and location of landmarks.[51] Cardillo and Sid-Ahmed reported a mean recognition rate of 75% within 2mm, however, there was a steady improvement in the performance of the software program throughout the experiment. The performance increased from 60% recognition rate to 85% recognition rate from the start of training the algorithm to the end.[51]

Rudolph et al. (1998) used a spatial spectroscopy-based detection algorithm that involved statistical pattern recognition.[52] They reported a low rate of landmark detection accuracy, with only 13% of landmarks successfully identified within 2mm and a mean error of 3.07mm with wide standard deviation values.

The first study to attempt automatic landmark detection with active shape models (ASM) was carried out by Hutton et al. (2000).[36] The reported accuracy was low at only 35% of landmarks successfully identified within 2mm and a mean error of 5.2mm. ASM was found to not yet be suitable for clinical use. Saad et al. (2005) extended the work of Hutton et al (2000) and used an active appearance model (AAM) approach which resulted in greater accuracy than the ASM method with 61% of landmarks successfully identified within 2mm and a mean error of 3.12mm.[53] and the authors also reported that AAM had a large computer memory requirement, which would be a further limitation for use. Vucinic et al. (2010) used a multi-resolution approach to AAM which resulted in a similar accuracy of 61% within 2mm but a reduced mean error of 1.68mm.[19]

These studies demonstrated continuously improving results in terms of increased accuracy over previously tested methods, the ability to identify more complex landmarks and the successful use of lower quality radiographs. However, despite these improvements, the results with model-based approaches were found to not be reliable enough for clinical use.

Table 1.2 Summary of the accuracy reported by studies using model-based approaches

Author (Year)	No. of test X-rays	No. of landmarks and accuracy	Techniques	Comment
Cardillo (1994)[51]	40	20 landmarks 75% < 2mm Mean error = not reported	Pattern matching	85% recognition rate (after 40 x-rays) Important study – used as a comparison in multiple future papers
Rudolph (1998)[52]	14	15 landmarks 13% < 2mm Mean error = 3.07mm	Spatial spectroscopy Statistical pattern recognition	Low resolution images used The authors suggested that if image resolution improved so would the accuracy of landmark identification – not demonstrated No statistical difference between manual and automatic methods
Hutton (2000)[36]	63	16 landmarks 35% < 2mm (13% < 1mm) Mean error = 4.08mm	Active shape models (ASM)	Poor quality images excluded Porion was the least accurately identified landmark (7.3mm error) B Point was the most accurately identified landmark (2.6mm error)
Saad (2005)[53]	7	18 landmarks Mean error = 3.1mm	Active appearance model	More accurate than ASM 27 pre-treatment lateral cephalograms randomly selected 20 used for building the active appearance model 7 used for testing
Vucinic (2010)[19]	60	17 landmarks 61%<2mm Mean error = 1.68mm	A multi-resolution active appearance model	Demonstrated increased overall precision and detection of landmarks on poorer quality radiographs than previous studies

1.13 Soft Computing Approaches

Soft computing approaches are a multidisciplinary field that comprise several different techniques. The aim of soft-computing approaches is the development of intelligent machines that can be programmed to think like humans and that have the ability to solve complex, non-linear problems[13]. Various soft computing disciplines have been used in the field of automatic landmark detection including a pulse coupled neural network, a support vector machine, a genetic programming algorithm and fuzzy neural networks. Table 1.3 summarises studies that examined the ability of soft-computing approaches to successfully locate anatomical landmarks on lateral cephalograms. These studies were conducted in the early 2000s.

Innes et al. (2002) used a pulse coupled neural network to locate three landmarks; the tip of the nose, B point and Sella, on 109 lateral cephalograms. The tip of the nose and B point were located with a high level of accuracy at 93.6% and 88.1% respectively while Sella had a much lower successful detection rate at 36.7%.[54] The mean error was not reported. The limitations of this study were that the parameters of the pulse coupled neural network had to be set manually and that only three landmarks were examined.

A support vector machine is a computer algorithm that can be trained to solve classification or regression problems.[55] Chakrabartty et al. (2003) used a support vector machine to locate 8 landmarks on 40 lateral cephalograms and reported a high successful detection rate of 93.87% of the landmarks located within 1mm.[56]

Ciesielski (2003) used genetic programming to develop a landmark detection algorithm for automated cephalometric landmark detection.[57] This involved the identification of a specific square-shaped search area around each landmark to be identified. Four landmarks were included in the study: pronasale, labrale superius, upper incisor tip and sella. Detection rates ranged from 63.9% (30.6% false alarm rate) to 100% (0% false alarm rate) with an average successful detection rate of 85% within 2mm.[57] A limitation of this study is that only 4 landmarks were examined.

Computers are generally programmed to provide clear yes or no answers to problems but, complex problems, particularly in the field of biology or medicine, often cannot be solved by simple yes or no answers. The human brain is capable of more nuanced and complex functions such as recognition, understanding and learning, which can assess a problem or question and develop an answer. The concept of fuzzy logic was introduced by Lofti Zadeh[58] to attempt to "make computers think like people".[59] The name fuzzy comes from the fact that the boundaries for solving complex problems are blurred or fuzzy rather than a crisp delineation between yes or no. A fuzzy neural network is a series of algorithms that can be used to identify the parameters of a fuzzy set, which is a set with fuzzy or gradual boundaries.[59] El-Feghi et al. (2004) used a fuzzy neural network with a template matching technique to detect 20 landmarks on 565 radiographs with a high level of success. They reported that 90% of the landmarks were successfully detected within 2mm.[60] This study had a very large sample size and this technique demonstrated a higher success rate than previously described techniques.

Overall, soft computing methods demonstrated improvement in accuracy but due to their limitations for clinical use researchers began developing hybrid algorithms to attempt to increase the accuracy and reliability of automated localisation of landmarks.

Table 1.3 Summary of the accuracy reported by studies using soft computing approaches

Author (Year)	No. of test X-rays	No. of landmarks and accuracy	Techniques	Comment
Innes (2002)[54]	109	3 landmarks 72% < 2mm Mean error = not reported	PCNN (pulse coupled neural networks)	93.6% soft tissue recognition rate (tip of nose) 88.1% hard tissue recognition rate (B-Point) Sella (S) very low success rate (36.7%) Only 3 landmarks
Chakrabartty (2003)[56]	40	8 landmarks 93% < 1mm Mean error = not reported	Support vector machine	High accuracy reported Only 8 landmarks
Ciesielski (2003)[57]	36	4 landmarks 85% < 2mm	Genetic algorithms	Only 4 landmarks
El-Feghi (2004)[60]	565	20 landmarks 90% < 2mm Mean error = not reported	Fuzzy neural networks	Large sample size Method showed improvement in accuracy over previous work

1.14 Hybrid Approaches

Research in the field of automated cephalometric analysis evolved to the development of hybrid approaches to locate anatomical landmarks. This involved combining existing techniques with novel approaches to try to improve the accuracy of the automated systems. Table 1.4 summarises studies that examined the ability of various hybrid approaches to automatically localise landmarks on lateral cephalograms. These studies were carried out and reported between the early 2000s and the mid-2010s.

The most frequently used technique in these studies was edge detection, which was combined with several other techniques to create hybrid approaches. Edge detection was used in combination with model-based approaches,[22] pattern matching,[61] active shape models,[62] probabilistic relaxation,[63] template matching,[13, 64] and an affine transform matrix, which is a method of relating landmarks on a reference image to a target image.[65]

Neural networks were combined with knowledge-based approaches[66] and active shape models.[67] Rueda et al. (2006) combined an active appearance model with a top-hat transformation approach, which is a method of improving the quality of the x-ray images and making the sample more uniform and reducing the variability of the x-rays.[68] All of these methods had varying degrees of success. Kaur et al. (2013) combined Zernike moments-based feature matching with template matching.[69] Zernike moments are used to improve the quality of an image and to get an approximate landmark location followed by template matching to narrow down the location of the landmark of interest.[69]

These hybrid approach studies reported varying degrees of success. Five of them reported low levels of accuracy ranging from 23-66.66% of landmarks successfully located within 2mm.[13, 22, 67-69]

Three studies reported high levels of accuracy ranging from 80-93% of landmarks successfully located within 2mm,[64-66] however the methodologies were limited by small sample size[64] or a small number of landmarks being assessed[65, 66].

A study by Grau et al. (2001) found very promising results using a combination of knowledge-based edge detection and template matching.[61] These methods were both used individually in previous studies and combined for this study.[47, 51] The study attempted to locate 17 landmarks on 20 radiographs and the reported overall success rate was 90.3%, with 88.6% of landmarks identified correctly within a 2mm precision range.[61] This might be suitable for clinical use if it facilitated manual editing of landmarks.

Overall, the reported hybrid approach techniques could not be recommended for widespread use due to either the accuracy of automated landmark identification being too low or due to limitations of the studies that reported high accuracy rates.

Table 1.4 Summary of the accuracy reported by studies using hybrid approaches

Author (Year)	No. of test X-rays	No. of landmarks and accuracy	Techniques	Comment
Liu (2000)[22]	10	13 landmarks 23% < 2mm (8% < 1mm) Mean error = 2.86mm	Edge detection Model fitting Knowledge-based algorithms	All errors made by automatic identification were larger than by manual identification Sella = most accurate Orbitale = least accurate
Grau (2001)[61]	20	17 landmarks 90.3% < 2mm Mean error = 1.1mm	Edge detection Pattern matching	Combination of edge detection and pattern matching in this study was more successful than each method used individually in previous studies Manual step involved for line selection
Yang (2001)[64]	11	16 landmarks 80% < 2mm Mean error for linear measurements < 0.9mm Mean error for angular measurements < 1.2°	Filtering Knowledge-based edge tracing Changeable templates	Small sample size Not all landmarks identified automatically
Giordano (2005)[66]	26	8 landmarks 85% < 2mm (73% < 1mm) Mean error = 1.07mm	Cellular neural networks Knowledge-based extraction	Only 8 landmarks
Yue (2006)[62]	86	12 landmarks 71% < 2mm	Filtering Edge tracking Pattern matching Active shape models	Modification on the original active shape models Sampling technique not discussed Active shape models not ideal for landmark detection

Table 1.4 Summary of the accuracy reported by studies using hybrid approaches (Continued)

Author (Year)	No. of test X-rays	No. of landmarks and accuracy	Techniques	Comment
Rueda (2006)[68]	96	28 landmarks 50.04% < 2mm Mean error = 2.48mm	Top hat transformation approach Active appearance model	Initial model used to extract the template is not explained
Mohseni (2007)[65]	40	8 landmarks 93% < 2mm	Affine transform matrix Edge detection	Only 8 landmarks
Kafieh (2007)[67]	63	16 landmarks 24% < 1mm 61% < 2mm 93% < 5mm	Learning vector quantization (LVQ) neural network Active shape model	Same dataset as Hutton et al.[36]
Favaedi (2010)[63]	4	21 landmarks 30.95% </= 1mm 71.4% </= 2mm	Probabilistic relaxation Edge-detection	Very small sample size
Shahidi (2013)[13]	40	16 landmarks 43.75% < 2mm (12.5% < 1mm) Mean Error = 2.59mm	Template-matching algorithm Edge-enhancement	Difficulty with ANS
Kaur (2013)[69]	85	18 landmarks 66.66% < 2mm Mean error = 1.84mm	Zernike moments based global feature matching Template-based local feature matching	Complex algorithm

1.15 Artificial Intelligence (AI) Based Models

1.15.1 Background

Artificial Intelligence (AI) is a general term used to describe a computer being able to think like a human and exhibit intelligent behaviour.[70] AI was first introduced in 1956 and is described as the science and engineering of building intelligent machines.[70, 71] These machines can be programmed to "learn" information and be trained to become more accurate over time. The development of AI has led to various applications across many areas of medicine including robotics, diagnostic tools, statistics, and biology.[70] Technological advancements in the field of AI have resulted in encouraging results in orthodontics including orthodontic diagnosis, treatment planning and clinical decision-making.[44, 46] The most commonly used application of AI in orthodontics is for identifying cephalometric landmarks.[46]

Various artificial intelligence (AI) techniques are used to automatically locate anatomical landmarks on lateral cephalograms. Machine learning is the term used to describe the development of computer programs and algorithms that automatically improve with experience.[72, 73] Using training data, machine learning algorithms "learn" to recognise patterns and be able to predict decisions without being specifically programmed to do so.[74] Deep learning is a subfield of machine learning and was developed to mimic the way the human brain works.[72] Deep learning algorithms are made up of artificial neural networks, which are computer systems built to mimic biological neural circuits.[73] These are arranged in layers and the word 'deep' is used to reflect the number of layers in the neural network. Convolutional neural networks are a type of deep neural network commonly used in the examination and analysis of images.[75] They are designed to mimic the biological visual cortex.[75] Random forests are a type of machine learning made up of multiple learning algorithms known as decision trees.[76] Decision trees are capable of predictive modelling by making observations about an item, a landmark for example, which is represented in the branches of the decision tree, and formulating conclusions about the location of that landmark, which is represented in the leaves of the decision tree.[77]

Cross-validation is a statistical method used to predict the success of machine learning models in practice by measuring how the results of a statistical analysis will generalise to an independent data set.[78] Several AI-based studies assessing the accuracy of AI-based computer programming techniques for automated cephalometric analysis use 4-fold cross validation to predict the performance of the technique or algorithm being assessed.

Over the past 10 years there have been substantial improvements in the accuracy of automatic cephalometric landmark detection as a result of advancements in artificial intelligence techniques.[79] Table 1.5 summarises studies that used various artificial intelligence-based models to assess the ability to detect anatomical landmarks successfully and accurately or to measure cephalometric angles on lateral cephalograms. These studies were conducted between 2009 and 2021.

1.15.2 Early AI Studies

The use of artificial intelligence for automated cephalometric analysis was first discussed in a significant way by Sommer et al. in 2009.[8] They compared traditional hand tracing to both semi-automatic and fully automatic cephalometric analysis performed using the orthodontic software Orthometric® in the measurement of 12 cephalometric angles on 72 lateral cephalograms.[8] They found that the measurements calculated using the semi-automatic mode were as accurate as hand tracing with all of the angles measured within the required 2° of the hand traced values. However, the measurements calculated using the fully automatic mode had multiple errors and none of the angles measured within the required 2° of the hand traced values. The authors recommended that if using the computer software Orthometric®, it would be essential to check all automatically located landmarks prior to performing an analysis and to correct their position where necessary.[8]

A study by Leonardi et al (2009) examined the ability of Cellular Neural Networks (CNNs) to accurately locate 10 landmarks on 41 lateral cephalograms and found a high level of accuracy with 100% of the automatically detected landmarks located within 0.59mm of the manually identified landmarks.[80] These results were very promising for the use of an automated landmark detection method however, two limitations of this study are that the 10 landmarks alone are not sufficient to perform a cephalometric analysis and also that training is required for each new landmark to be detected.[80]

1.15.3 Grand Challenge

With the development of artificial intelligence techniques to automatically identify anatomical landmarks, AI challenges were introduced at the International Symposium on Biomedical Imaging organised by the Institute of Electrical and Electronics Engineers (IEEE ISBI) in 2014.[40] Researchers competed by developing AI-based models to analyse a dataset of lateral cephalograms and were assessed on a number of parameters to determine the most accurate program.[40] The challenge involved identification of 19 specific anatomical landmarks.

The IEEE ISBI 2014 Grand Challenge in Cephalometric Image Analysis Dataset consists of 300 lateral cephalograms. This is divided into a 100-image training set, and two testing datasets each consisting of 100 images.[40] The IEEE ISBI 2015 Grand Challenge in Cephalometric Image Analysis Dataset consists of 400 lateral cephalograms. This is divided into a 150-image training set, 140 images for training and 10 for validation and then two testing datasets. The Test 1 dataset has 150 images, and the Test 2 dataset has 100 images.[81] Each of the lateral cephalograms in these datasets consists of the original x-ray image and two sets of manually located landmarks by two specialist orthodontists, Expert 1 had 15 years' experience and Expert 2 had 6 years' experience.[40] The ground truth for each landmark is described as the mean position between the two points located by the two clinical experts. The mean intra-examiner variability for Expert 1 and Expert 2 was 0.9mm and 1.73mm respectively.[40]

The 2014 IEEE ISBI grand challenge in cephalometric image analysis was won by Ibragimov et al. (2014), a team of researchers from Slovenia, who used game theory and random forests (RF) to analyse the dataset and found that 71.48% of the 19 landmarks were located within 2mm with a mean error of 1.92 ± 1.88mm.[82]

The 2015 IEEE ISBI grand Challenge in cephalometric image analysis was won by Linder and Cootes (2015) from Manchester in the United Kingdom, who used Random Forest regression voting combined with a model-based approach to analyse the dataset and found that 74.95% of the 19 landmarks were located within 2mm with a mean error of 1.67 ± 1.65mm.[41]

The same researchers teamed up with another group of researchers from Taiwan and conducted a study in 2016, which built on the Random Forest regression voting method used in 2015 to develop a Fully Automatic Landmark Annotation (FALA) system and examined its accuracy using 2015 IEEE ISBI grand challenge dataset.[83] They found that 84.5% of the 19 landmarks were located within 2mm with a mean error of 1.2mm.[83] The authors reported that accuracy of landmark identification was found to be within the inter-observer variability between two clinical experts.[83]

Most of the other studies relating to examining the accuracy of AI-based models in landmark detection also used the IEEE ISBI 2014 and 2015 grand challenges in cephalometric image analysis x-ray dataset,[79, 84-95] and therefore their results are comparable. Two tests are performed and the highest mean successful detection rates within 2mm ranged from 80.2% to 82.69%,[90-94] with the lowest mean error reported by Gilmour et al. (2020) as 1.17 ± 0.80mm.[93] These results are summarised in Table 1.6. For an individual test, four studies reported that more than 86% of the 19 landmarks were located within 2mm.[91-94] A 4-fold cross-validation to predict the success of the proposed method was reported in three studies and resulted in higher successful detection rates (SDR) <2mm. Lindner et al (2016) reported an SDR <2mm of 84.7% with a mean error of 1.2 ± 0.6mm.[83] Zhong et al (2020) reported an SDR <2mm of 86.06% with a mean error of 1.22 ± 2.45mm.[92] Gilmour et al (2020) reported an SDR <2mm of 86.72% with a mean error of 1.07 ± 0.95mm.[93] These results are summarised in Table 1.6.

1.16 Recent Artificial Intelligence Studies

The majority of recent studies examining the success of AI-based models use techniques that fall into the categories of deep learning and artificial neural networks,[39, 79, 85, 87-98] or random forests.[41, 42, 82-84, 86] Park et al (2019) and Hwang et al (2020) published two papers related to a novel deep learning method entitled You-Only-Look-Once version 3 (YOLOv3).[39, 96] In the first study, YOLOv3 was compared to another deep-learning method called Single Shot Multibox Detector (SSD). Both methods were trained using 1028 cephalometric radiographs and then tested with a new sample of 283 radiographs on their ability to identify 80 landmarks. The results found that YOLOv3 was more successful than SSD in terms of accuracy and mean computational time and had a landmark successful detection rate (SDR) <2mm of 80.4%.[39] In the second study, YOLOv3 was compared to human experts using the same methodology and sample size as the previous study. The 80 landmarks were identified by YOLOv3 and the human examiners twice. The results found that the mean error between YOLOv3 and human experts was 1.46 ± 2.97mm. This was higher than the mean error between human examiners which was 1.5 ± 1.48mm. However, YOLOv3 consistently identified landmarks in identical positions, eliminating intra-examiner variability as it suggested that AI-based software can be more reliable than human experts.[96]

1.16.1 Commercially Available Software

Anuwongnukroh et al. (2017) conducted a study to assess the accuracy of the Carestream Dental automatic cephalometric analysis software program (Version 6.14) in identifying 13 landmarks on 30 radiographs compared to traditional hand tracing.[43] Each manually and automatically identified landmark was measured on an x and y axis created on each lateral ceph by passing the x and y axis lines through the centre of the ear rod. The accuracy was examined by measuring the mean differences between the x coordinates for each landmark and the y coordinates for each landmark and a one-sample t test was used to analyse the differences.[43] This method allowed assessment and quantification of the direction of error. Statistically significant mean differences ($p<0.05$) were found for five landmarks (Or, A Point, Me, LIT and LIA) in the horizontal direction and for seven landmarks (Or, A Point, UIT, UIA, B Point, Me and LIA) in the vertical direction.[43] There were no statistically significant differences between the automated and the manual tracing in five of the landmarks (S, N, Gn, Pog and Go).[43] The successful detection rate within 2mm was not measured, nor was the mean radial error which makes it more difficult to compare with other methods. The authors concluded that the Carestream Dental automatic cephalometric analysis software program (Version 6.14) was unable to compete with the traditional manual tracing method in terms of accuracy of landmark detection. They recommended that if using the program for cephalometric analysis, the orthodontist should correct or modify the landmark positions before carrying out an analysis.[43]

The same authors conducted another study examining the accuracy of the Carestream Dental automatic cephalometric analysis software (Version 6.14) in recording cephalometric angular and linear measurements compared to traditional hand tracing.[99] They examined eight linear and nine angular measurements. They found that 13 out of the 17 measurements performed automatically showed statistically significant differences compared to the manual method (76.47%). Only four out of the 17 measurements did not show any statistically significant differences (23.53%). The authors concluded that the cause of the low accuracy was the inability of the software program to accurately locate anatomical landmarks and therefore Carestream Dental automatic cephalometric analysis software (Version 6.14) is not accurate enough for clinical use unless the automatically detected positions of the landmarks are adjusted or corrected prior to performing an analysis.[99]

Jeon et al. (2021) carried out a study comparing the results of an AI-based automated cephalometric analysis software programme, CephX™ (ORCA Dental AI Inc., Herzliya, Israel) with a conventional cephalometric software programme on a sample of 35 radiographs.[100] Twenty-six cephalometric measurements were recorded using both methods: 15 skeletal, 9 dental and 2 soft tissue. Statistically significant differences were found between the two methods for three of the measurements: saddle angle, upper incisor to N-A line and lower incisor to N-B line.[100] Overall, they found that the dental measurements had more variability than the skeletal or soft tissue measurements. The authors concluded that the CephX™ software programme may be suitable for clinical use. The accuracy of landmark identification was not examined in this study. CephX™ is the company that Planmeca Romexis® have partnered with for their automated cephalometric analysis service.[101]

Meriç and Naoumova (2020) carried out a study comparing the accuracy of fully-automated, semi-automatic, app-aided and manual cephalometric analysis using a sample of 40 pre-treatment lateral cephalograms.[102] The fully-automated programme was CephX™ (ORCA Dental AI, Las Vegas, NV), the semi-automatic was Dolphin Imaging® 13.01 (Dolphin Imaging and Management Solutions, Chatsworth, Calif, USA), the app-aided software was the CephNinja® 3.51 App (Cyncronus LLC, WA, USA) and the hand tracings were carried out by three operators and a manual ground truth was obtained using the mean values of the three operators.[102] The authors used CephX™ with and without manual correction of landmarks. They measured the accuracy of eight angular and four linear parameters measured on each radiograph. Their findings were that CephX™ with manual corrections, Dolphin Imaging® and CephNinja® App were comparable to hand tracings but that CephX™ without manual corrections was the least accurate method and needs to be improved for clinical use.[102] The time for analysis was shortest with CephX™ without manual corrections followed by CephX™ with manual corrections.[102]

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Sommer (2009)[8]	72	N/A	12 angles Semi-automatic: All angles < 2 degrees except for NLA Automatic: All angles > 2 degrees (outside tolerance limit)	Orthometric cephalometric analysis software Comparing hand tracing to semi-automatic and fully-automatic cephalometric analysis	Semi-automatic = angles measured with the same reliability as hand tracing Fully-automatic = multiple errors. Mandatory to check all automatically detected landmarks and to correct them if required prior to analysis
Leonardi (2009)[80]	41	N/A	10 landmarks 100% detection <0.59mm	Cellular Neural Network (CNN)	High accuracy Training required for each new landmark to be detected
Chu (2014)[84]	100	100	19 landmarks 41.93% < 2mm	Random Forest (RF) regression-based landmarks detection combined with a model-based approach	IBSI 2014 Grand Challenge in Dental X-Ray Imaging Dataset Very low accuracy
Ibragimov et al (2014)[82]	200	100	19 landmarks 71.48%% < 2mm Mean error = 1.92 ± 1.88mm	Game theory and Random Forests (RF)	Winner of the 2014 IEEE ISBI Cephalometric X-Ray Image Analysis Challenge
Ibragimov et al (2015)[42]	250	150	19 landmarks 71.7% < 2mm Mean error = 1.84 ± 1.79mm	Random Forest (RF) regression-based landmarks detection combined with a model-based approach	Runner Up of the 2015 IEEE ISBI Cephalometric X-Ray Image Analysis Challenge

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Lindner and Cootes (2015)[41]	250	150	19 landmarks 74.95% < 2mm Mean error = 1.67 ± 1.65mm	Random Forest (RF) regression-based landmarks detection combined with a model-based approach	Winner of the 20115 IEEE ISBI Cephalometric X-Ray Image Analysis Challenge
Lindner (2016)[83]	250	150	19 landmarks. 4-fold cross validation: 84.5% <2mm 92.6% <3mm 96.3% <4mm Mean error = 1.2 ± 0.6mm	Fully Automatic Landmark Annotation (FALA) System	Accuracy of automatic landmark identification was within the inter-observer variability between 2 clinical experts – 83.5% average classification accuracy
Arik (2017)[85]	250	150	~1-2% higher success detection rate for a 2mm range compared with other leading techniques in the field. ~76% average classification accuracy for test set	Convolutional Neural Networks (CNNs) Cephalometric x-ray image dataset was used to train CNNs for recognition of landmark patterns	Complicated paper Promising results Authors predict improvement of accuracy with time as further training is undertaken
Anuwongnukroh (2017)[43]	30	N/A	13 landmarks Horizontal and vertical directions 5 (38.46%) landmarks not accurate in horizontal direction 7 (53.85%) landmarks not accurate in vertical direction 4 landmarks not accurate in either direction	Carestream Dental Cephalometric Analysis Software, version 6.14	Examiner not calibrated. A-point in the horizontal direction and LIA in the vertical direction were the least accurately identified points. Carestream Dental, version 6.14 found to be unable to compete with human expert in terms of accuracy of landmark identification

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Anuwongnukroh (2018)[99]	108	N/A	8 linear and 9 angular measurements. 23.53% (4/17) of the automatic measurements showed no statistically significant differences compared to the manual method	Carestream Dental Cephalometric Analysis Software, Version 6.14 compared to manual measurement with protractor	Carestream Dental, version 6.14 not as reliable as manual analysis Software program could not interpret negative results therefore sample size for automatic analysis reduced and varied
Wang (2018)[86]	300 165	150	Dataset 1 = 19 landmarks Dataset 2 = 45 landmarks Dataset 1 = 73.37% < 2mm Dataset 2 = 72.08% < 2mm Dataset 1 = Mean error 1.69 ± 1.43mm Dataset 2 = Mean error 1.71 ± 1.39mm	Multiresolution decision tree regression voting	Two datasets: 1. IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset 2. Dataset from Peking University
Nishimoto (2019)[98]	153	66	10 landmarks Mean error reported in pixels 17.02 ± 11.13 pixels	Customised Convolutional Neural Network	Lateral cephalograms collected on the internet
Park (2019)[39]	283	1028?	80 landmarks 80.4% < 2mm YOLOv3 outperformed SSD in accuracy and time for analysis YOLOv3 demonstrated a 5% higher accuracy compared with previously published studies	You-Only-Look-Once version 3 (YOLOv3) – Deep learning method	Promising method. Human Expert identified gold standard. YOLOv3 compared to Single Shot Multibox (SSD) method.

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Zhong (2019)[92]	150	150	19 landmarks 81.46% < 2mm (average of 2 tests) Mean error 1.14 ± 1.03mm 4-fold cross-validation: 86.06% <2mm Mean error 1.22 ± 2.45mm	Deep regression model	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset
Qian (2019)[90]	250	150	19 landmarks 82.5% < 2mm 81.8% < 2mm	Convolutional Neural Network	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset
Dai (2019)[88]	200	100	19 landmarks 35-40% < 2mm for each landmark Mean error = 2.5-7.5mm for each landmark	Adversarial encoder-decoder networks	IBSI 2014 Grand Challenge in Dental X-Ray Imaging Dataset Low accuracy and large mean error
Goutham (2019)[89]	250	150	7 landmarks 65.13% < 2mm	Modified u-net	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset Small number of landmarks examined Low accuracy

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Chen (2019)[87]	250 1857	150	19 landmarks Dataset 1, Test 1 = 86.67% < 2mm Dataset 1, Test 2 = 75.05% < 2mm Best result on other datasets = 94.73% < 2mm Dataset 1, Test 1 = Mean error 1.25mm Dataset 1, Test 2 = Mean error 1.47mm Best results on other datasets = Mean error 0.88mm.	End-to-end deep learning framework	Two dataset categories: 1. IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset 2. 5 new datasets collected (A-E) for the purpose of the study Technique performs well on unseen datasets
Song (2020)[91]	250 100	150	19 landmarks Dataset 1, Test 1 = 86.4% < 2mm Dataset 1, Test 2 = 74.0% < 2mm Average = 80.2%<2mm Dataset 2 = 62% < 2mm Dataset 1, Test 1 = Mean error 1.077mm Dataset 1, Test 2 = Mean error 1.542mm Dataset 2 = Mean error 2.1mm	Image pre-processing to detect region of interest (ROI) Convolutional Neural Network – ResNet50	Two datasets: 1. IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset 2. Dataset from Shandong University High computational time Lower results when tested on dataset that was not used for training

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Oh (2020)[94]	250	150	19 landmarks Test 1 = 86.2% <2mm Test 2 = 75.89% <2mm Average = 81.05% <2mm Mean error reported in pixels Test 1 = 11.77 ± 10.13 pixels Test 2 = 14.55 ± 8.22 pixels	Convolutional Neural Network	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset
Kunz (2020)[97]	50	1792	18 landmarks Absolute mean differences >0.37° for angular measurements >0.2mm for metric parameters >0.25% for the proportional parameter No statistically significant differences (except for SN-MeGo)	Customised Convolutional Neural Network	Did not report successful detection rate (SDR) <2mm

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Gilmour (2020)[93]	250	150	19 landmarks Test 1 = 88.32% < 2mm Test 2 = 77.05% < 2mm Average = 82.69% <2mm Mean error Test 1 = 1.01 ± 0.85mm Test 2 = 1.33 ± 0.74mm Average = 1.17± 0.95mm 4-fold cross-validation: 86.72% <2mm Mean error = 1.07 ± 0.95	Modified ResNet34 combined with a custom image pyramids approach	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset Lowest mean error ± SD reported in the literature (1.17±0.95mm, average between Test 1 and Test 2)
Hwang (2020)[96]	283	1028	80 landmarks Mean detection error between AI and human was 1.46 ± 2.97mm Mean difference between human examiners was 1.59 ± 1.48mm Comparison of detection errors between AI and human examiners <0.9mm – not clinically significant	You-Only-Look-Once version 3 (YOLOv3) – Deep learning method	AI method was as accurate as human expert AI method was more consistently accurate than human expert AI repeated detection error = 0mm Human intra-examiner variability of repeated manual detections: detection error = 0.97 ± 1.03mm

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Kim (2020)[95]	400	1675	23 landmarks Dataset 1: 82.92% <2mm Mean error = 1.37mm 91.17% classification accuracy Dataset 2: 84.53% < 2mm Mean error 1.16mm 83.13% classification accuracy	Two-stage automated algorithm with a stacked hourglass deep learning model	A web-based application was built by the authors using the proposed algorithm
Lee (2020)[79]	250	150	Mean landmark error = 1.53 ± 1.74mm 82.11% <2mm 92.28% <3mm 95.95% <4mm	Bayesian Convolutional Neural Networks (BCNN)	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset
Noothout (2020)[103]	250	150	19 landmarks Test 1 Error Range = 0.46-2.12mm Test 2 Error Range = 0.42-4.32mm ~78% <2mm ~90% <3mm ~95% <4mm	Fully Convolutional Neural Network (FCNN)	IBSI 2015 Grand Challenge in Dental X-Ray Imaging Dataset Exact successful detection rates (SDR) for specified ranges not given – results given on bar chart compared to other studies

Table 1.5 Summary of the accuracy reported by studies using artificial intelligence (AI) based models (Continued)

Author (Year)	No. of test X-rays	No. of training X-rays	No. of landmarks/angles and accuracy	Techniques	Comment
Meriç and Naoumova (2020)[102]	40	N/A	8 angular and 4 linear measurements CephNinja® and Dolphin® were comparable to manual tracings CephX™ with correction of landmarks was comparable to CephNinja® and Dolphin CephX™ without correction of landmarks least accurate	Dolphin Imaging® 13.01 (computerised tracing) CephNinja® 3.51 App (App-aided tracing) CephX™ (Web-based fully automated tracing) – with and without correction of landmarks	Compared to manual tracing by 3 operators – "manual ground truth" Shortest analysing time with CephX™
Jeon (2021)[100]	35	N/A	26 angular and linear measurements 3 not accurately identified Dental measurements less precise than skeletal measurements	CephX™ software programme Convolutional neural network	Did not measure accuracy of landmark identification

Table 1.6 The ten highest reported Successful Detection Rates (SDR) for a 2.0mm precision range in the literature to date

Author (Year)	Successful Detection Rate (SDR) <2mm (%)	Average Mean Radial Error (MRE±SD) (mm)
Song (2020)[91]	80.20 (average of 2 tests)	1.31 (SD not reported)
Park (2019)[39]	80.40	Not reported
Chen (2019)[87]	80.86 (average of 2 tests)	1.33 (SD not reported)
Oh (2020)[94]	81.05 (average of 2 tests)	Not reported in mm
Zhong (2019)[92]	81.46 (average of 2 tests) 86.06 (4-fold cross-validation)	1.27 ± 0.86 1.22 ± 2.45
Lee (2020)[79]	82.11	1.53 ± 1.74
Qian (2019)[90]	82.15 (average of 2 tests)	Not reported
Gilmour (2020)[93]	82.69 (average of 2 tests) 86.72 (4-fold cross-validation)	1.17 ± 0.80 1.07 ± 0.95
Kim (2020)[95]	83.73 (average of 2 tests)	1.27 ± 0.93
Lindner (2016)[83]	84.50	1.20 ± 0.60

1.17 Planmeca Romexis® Dental Imaging Software

Planmeca Romexis® (Version 5.1.1.R) dental imaging software offers a package for automated cephalometric tracing and analysis after partnering with CephX Technologies Limited. Planmeca Romexis® utilises CephX's intelligent AlgoCeph algorithm, an artificial intelligence method, which uses machine learning technology.[101]

1.18 Gap in the literature

There is limited literature available on the accuracy of commercially available automated cephalometric software programmes and no available literature on the accuracy of automated cephalometric analysis with Planmeca® Romexis (5.1.1.R) dental imaging software.

Chapter 2 Study Aim and Objectives

2.1 Study Aim

The aim of this study was:

- To determine if automated computerised cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software is as accurate as traditional hand tracing in terms of landmark identification, angular and linear measurements.

2.2 PICO

Population:	Orthodontic patients requiring a lateral cephalogram radiograph prior to treatment
Intervention:	Automated computerised cephalometric analysis
Comparison:	Traditional hand tracing cephalometric analysis
Outcome:	Accuracy of automated computerised cephalometric analysis relative to traditional hand tracing

2.3 Study Objectives

The objectives of this study are:

- To measure the level of inter-rater agreement of cephalometric analysis by hand tracing of two Consultant Orthodontists and the primary investigator.
- To measure the level of agreement for cephalometric analysis by hand tracing of the primary investigator against themselves at two different time points.
- To measure the level of agreement on cephalometric analysis between automated computerised cephalometric analysis using Planmeca Romexis® dental imaging software and traditional hand tracing.

2.4 Null Hypothesis

- There is no difference in accuracy between automated computerised cephalometric analysis with Planmeca Romexis® dental imaging software and traditional hand tracing.

Chapter 3 Materials and Methods

3.1 Study Design

An observational, comparative, cross-sectional study was conducted to assess the accuracy of automated computerised cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared to traditional hand tracing cephalometric analysis.

3.2 Ethical Approval

Ethical approval for this study was sought from the Faculty of Health Sciences Research Ethics Committee of Trinity College Dublin and approval was granted. A copy of the letter of approval is included in Appendix 2.

3.3 Inclusion and Exclusion Criteria

The inclusion criteria were pre-treatment lateral cephalograms with the teeth in occlusion on the cephalogram. The exclusion criteria were unerupted or missing incisors, unerupted teeth overlying the apices of the incisors, patients with obvious craniofacial deformity or asymmetry and diagnostically unacceptable image quality.

3.4 Sample Size

The sample sizes for this study were calculated based on the use of intraclass correlation coefficients (ICC) as the statistical method of assessing level of agreement. This calculation method determines the minimum sample size required for the desired effect size of ICC.[104] The calculation is taken from the Power Analysis and Sample Size (PASS) software where the alpha is set at 0.05 and the power is greater than 80%.[104, 105] The two key parameters in these sample size calculations are the value of ICC and the number of observations recorded by each rater.[104]

3.4.1 Inter-Examiner Reliability

To determine if the inter-rater agreement for cephalometric analysis by hand tracing of two Consultant Orthodontists and the primary investigator could be as high as 0.7 or greater, a minimum sample size of eight lateral cephalograms was required to achieve statistical significance for an alpha of 0.05 and a power of 90%.

3.4.2 Intra-Examiner Reliability

To determine the level of agreement for cephalometric analysis by hand tracing of the primary investigator against herself at two different time points, a minimum sample size of 13 lateral cephalograms was required for an alpha of 0.05 and a power of 90%.

3.4.3 Study

A sample size of at least 68 radiographs was required to achieve statistical significance for an alpha of 0.05 and a power of more than 90%.

3.5 Recruitment of Patients

Orthodontic patients, and parents or legal guardians of patients under 16 years of age, attending the Health Service Executive (HSE) Orthodontic Unit, SIMMS Building, Tallaght, requiring a pre-treatment lateral cephalogram were given a patient information leaflet regarding the study and invited by the study gatekeeper, the unit senior dental nurse, to participate. As this project involved the recruitment of children, a separate patient information leaflet written in child-friendly language was included with each information pack to ensure that these participants were as fully informed as possible about the nature of their involvement. The patients and parents were given time to read the patient information leaflet(s) and an opportunity to ask question prior to consenting to participate. The patient information leaflets explained that the participant's radiograph would be irreversibly anonymised prior to being measured and that, apart from the consent forms, no personal data would be collected or stored as part of the study. The adult participation information leaflet is included in Appendix 3, the parent/guardian participant information leaflet is included in Appendix 4 and the child participation information leaflet is included in Appendix 5. If the patients and parents agreed to participate, written consent was obtained from them. The adult consent form is included in the Appendix 6 and the parent/guardian consent form is included in Appendix 7. Recruitment ceased once 85 radiographs had been collected, as this was sufficient to meet the sample numbers required. The study was carried out prospectively as, at the time of commencement of this project, Planmeca Romexis® dental imaging software had recently been installed in the department and also to ensure that all the radiographs used in the study were taken by the same radiographer using the same cephalostat.

3.6 Anonymisation

The radiographs were exported from Planmeca Romexis® anonymously and saved as individual images in a folder by the study gatekeeper, thus irreversibly anonymising each radiograph. No key was created to enable identification of the patients once the radiograph had been anonymised. Suitability for selection was based on the inclusion and exclusion criteria above. Four radiographs were excluded, two as the patients were not in occlusion and two as the patients had unerupted teeth overlying the apices of the upper incisors. Thirteen radiographs were used to determine inter-examiner and intra-examiner reliability. These were not included in the study sample. The final sample size for the study was 68. The anonymised radiographs were imported into Planmeca Romexis® under a newly created "CM Study" patient chart and saved as numbers 1-68.

3.7 Type of Cephalometric Analysis

Planmeca Romexis® (Version 5.1.1.R) dental imaging software has the facility to select the type of analysis that the operator wishes to carry out on a particular radiograph. The two most frequently used analyses on the programme are the Eastman analysis and the Downs analysis and therefore these were selected to be used in the study. The Eastman analysis involved the identification of 13 anatomical landmarks, seven angular measurements and two linear measurements. The Downs analysis involved the identification of 14 anatomical landmarks, nine angular and one linear measurement. Two angular measurements, AB plane and cant of occlusal plane to Frankfort plane were excluded from the Downs analysis as they were difficult to ascertain on the tracing as a result of several overlapping lines due to the volume of information required for both analyses. There was some repetition and overlap of the two analyses and therefore, in combination, the final analysis used in the study included 16 anatomical landmarks, 12 angular and 3 linear measurements. These are described in detail in tables 3.1 and 3.2.

Table 3.1 Definition of landmarks identified in the study

Landmark	Abbreviation	Definition
Sella	S	The midpoint of the sella turcica (pituitary fossa)
Nasion	N	The most anterior point on the frontonasal suture in the midline
A-Point (subspinale)	A	The deepest point on the curved profile of the maxilla between the anterior nasal spine and the alveolar crest
B-Point (supramentale)	B	The deepest point on the curved profile of the mandibular symphysis
Anterior Nasal Spine	ANS	The tip of the bony anterior nasal spine in the midline
Posterior Nasal Spine	PNS	The tip of the bony posterior nasal spine in the midline
Upper Incisor Apex	UIA	The root apex of the most anterior maxillary central incisor
Upper Incisor Tip	UIT	The tip of the crown of the most anterior maxillary central incisor
Lower Incisor Apex	LIA	The root apex of the most anterior mandibular central incisor
Lower Incisor Tip	LIT	The tip of the crown of the most anterior mandibular central incisor
Pogonion	Pog	The most anterior point on the mandibular symphysis
Menton	Me	The most inferior part on the mandibular symphysis in the midline
Gonion	Go	The most posterior and inferior point on the angle of the mandible
Porion	Po	The upper- and outer-most point on the external auditory meatus
Orbitale	Or	The most inferior and anterior point on the infra-orbital rim
Gnathion	Gn	The most anterior and inferior point on the bony chin

Table 3.2 Definition of measurements used in the study

Measurement	Definition	Analysis
SNA	The angle formed between S, N and A	Eastman
SNB	The angle formed between S, N and B	Eastman
ANB	The angle formed between A, N and B	Eastman
UI Mx	The axial inclination of the maxillary incisors to the maxillary plane	Eastman
LI Md	The axial inclination of the mandibular incisors to the mandibular plane	Eastman/Downs
Interincisal (II) Angle	The angle formed by the intersection of the mandibular incisor axis and the maxillary incisor axis	Eastman/Downs
MMPA	The angle formed between the maxillary plane (ANS-PNS) and the mandibular plane (Go-Me)	Eastman
Overjet	Horizontal distance between the maxillary and mandibular central incisors	Eastman
LI/A-Pog	The linear distance from the tip of the lower incisor to the line from A-point to pogonion	Eastman
Facial Angle	The inferior internal angle between the facial plane (N-Pog) and the Frankfort plane	Downs
Angle of Convexity	The angle formed between the lines from N-A and A-Pog	Downs
FMPA	The angle formed between the Frankfort plane and the mandibular plane	Downs
Y-Axis	The angle formed between the Frankfort plane and S-Gn	Downs
LI-OcP	The axial inclination of the lower incisors to the occlusal plane	Downs
UI/A-Pog	The linear distance from the tip of the upper incisor to the line from A-point to Pogonion	Downs

3.8 Inter-Examiner Reliability

Inter-examiner reliability was examined by comparing the level of inter-rater agreement between two Consultant Orthodontists and the primary investigator. Three copies of eight lateral cephalograms were printed on plain radiographic films and distributed to each examiner with identical tracing kits containing 3H pencils, tracing paper sheets, a pencil sharpener, an eraser, an orthodontic protractor, descriptions of the two analyses, a list of anatomical landmarks, angular and linear measurements, and a printed template to fill in the results. The examiners were asked to hand trace each lateral cephalogram identifying specific anatomical landmarks and angular and linear measurements.

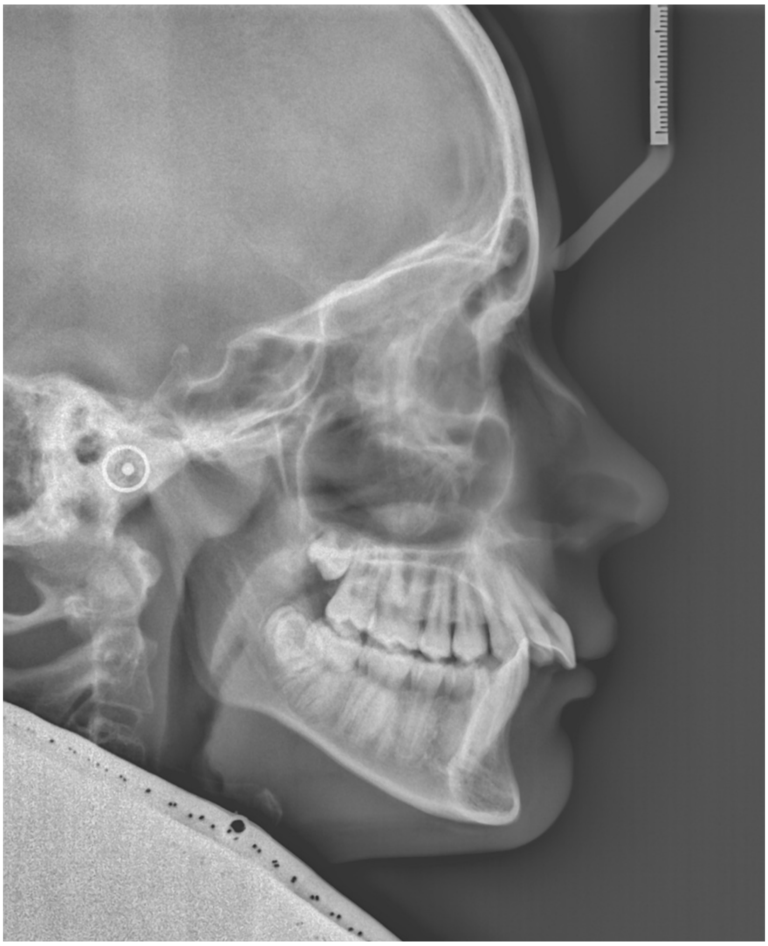
3.9 Intra-Examiner Reliability

Intra-examiner reliability was examined by comparing the level of agreement for cephalometric analysis by hand tracing of the primary investigator against herself at two different time points. Thirteen lateral cephalograms were used and this comprised the eight radiographs used for the inter-examiner reliability plus an additional five radiographs. The 13 radiographs were printed on plain radiographic films and traced by the primary investigator on two occasions, four weeks apart.

3.10 Hand Tracing

The radiographs were printed on laser films with an AGFA™ Drystar 5302 x-ray printer to facilitate hand tracing. Each printed radiograph was labelled as its corresponding number and the letter A, e.g., radiograph 1 was labelled “1A” to identify it as the original number 1 radiograph. An original radiograph from the study is shown in figure 3.1. Tracing was carried out with a sharp 3H pencil in a dark room with a light box. Tracing paper was placed over the lateral cephalogram allowing it to be lifted to check obscure landmarks. The nasion locator was traced on each radiograph to facilitate comparison with the automated analysis. The landmarks were labelled prior to analysis to minimise measurement error.[52] The angular and linear measurements for each radiograph were recorded in an Excel© (Microsoft Corporation) Workbook. As part of the D.Ch.Dent. (Orthodontics) programme, the lead examiner was trained in hand tracing cephalometric analysis. This training involved both lectures and hands-on practical sessions. The hand tracings and measurements for this study were conducted by the lead examiner in the same room, under the same conditions, with approximately six radiographs traced or measured per sitting to avoid examiner fatigue and to minimise errors.

Figure 3.1 A study radiograph



3.11 Automated Cephalometric Analysis

The automated cephalometric analysis was carried out using Planmeca Romexis® (Version 5.1.1.R) dental imaging software. Each study radiograph was analysed using the Eastman analysis and the Downs analysis. These have to be performed separately on each radiograph and two separate analyses are produced. The analysed radiographs were printed on plain radiographic films to facilitate comparison with the corresponding hand tracing. Each printed Eastman analysis radiograph was labelled as its corresponding number and the letter B, e.g., radiograph 1 was labelled “1B” to identify it as the Eastman analysis radiograph 1. A study radiograph with the automatically identified Eastman landmarks on it is shown is shown in figure 3.2. Each printed Downs analysis radiograph was labelled as its corresponding number and the letter C, e.g., radiograph 1 was labelled “1C” to identify it as the Downs analysis radiograph 1. A study radiograph with the automatically identified Downs landmarks on it is shown in figure 3.3. The angular and linear measurement values produced for each analysis were saved in an Excel© (Microsoft Corporation) Workbook.

Figure 3.2 A study radiograph with the Eastman analysis landmarks automatically identified

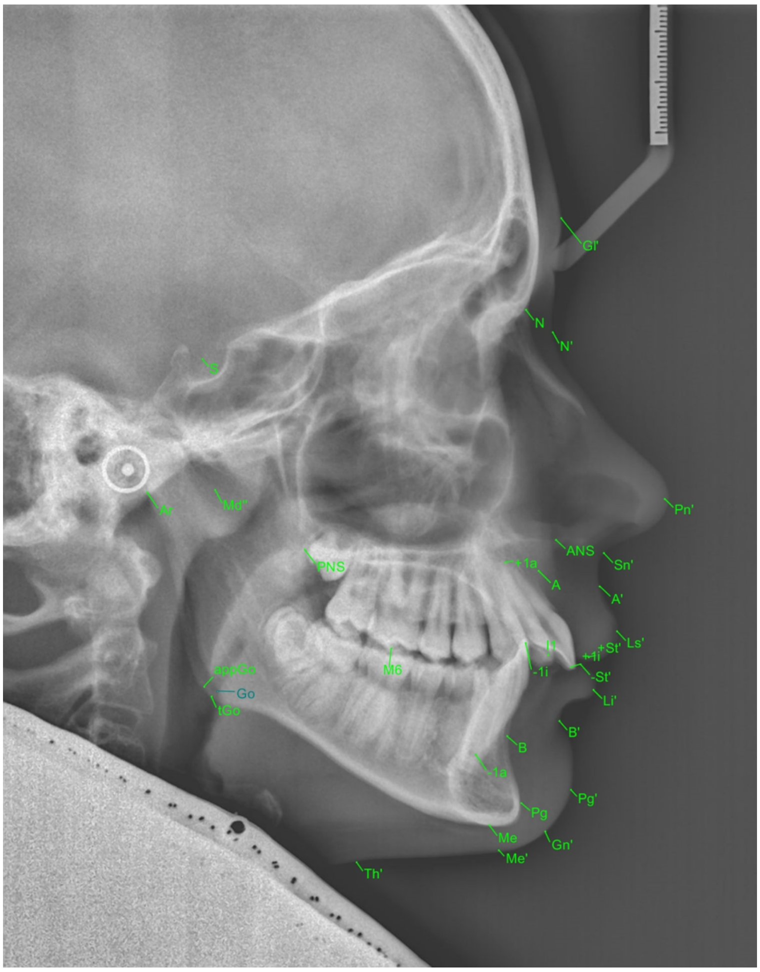
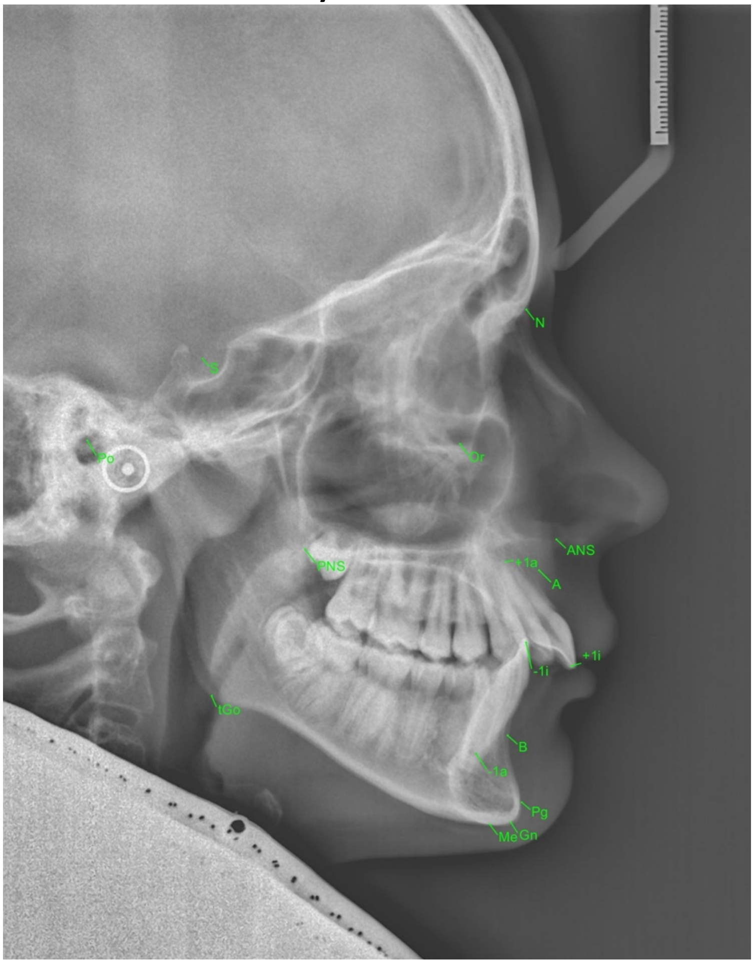


Figure 3.3 A study radiograph with the Downs analysis landmarks automatically identified



3.12 Landmark Identification Analysis

The hand tracing was oriented over the automated tracing for each radiograph using the Nasion locator and other anatomical areas such as Sella Turcica and the incisors. The landmarks were compared using 3 different methods:

1. 2mm precision range
2. Point-to-point error
3. X- and Y-Axis Coordinates

3.12.1 2.0mm Precision Range

The accuracy of the automated localisation of landmarks was assessed by determining if the automated landmark was identified within 2mm of the hand traced landmark. This was measured using a 4mm diameter circle with a crosshairs in the centre of it printed on acetate paper. The hand traced landmark was placed in the centre of the crosshairs and if the automated landmark lay within the circle it was recorded as Y and if the automated landmark lay outside the circle it was recorded as N in an Excel© (Microsoft Corporation) Workbook.

3.12.2 Point-to-Point Error

The point-to-point error or mean radial error was measured between the hand traced landmark and the automated landmark using digital callipers. The measurement was recorded for each landmark on each radiograph to two decimal places in an Excel© (Microsoft Corporation) Workbook.

3.12.3 X- and Y-Axis Coordinates

The inter-rater variability for anatomical landmark identification between two clinical experts has been reported in the literature to be approximately 85%.[83] This is described as two landmarks being identified within a 2mm precision range. In this study, any landmarks that were identified successfully by the automated system less than 85% of the time were further analysed by recording their x- and y-axis coordinates to determine the direction of error. This was done by creating an x- and y-axis on each tracing and recording an x- and y-axis measurement for each landmark. The y-axis was created by drawing a vertical line, parallel to the edge of the film through the centre of the ear rod on the radiograph. The x-axis was created by drawing a horizontal line perpendicular to the y-axis through the centre of the ear rod on the radiograph. The x- and y-axis values were measured using digital callipers. The hand traced value, the automated value, and the difference between them were recorded in an Excel© (Microsoft Corporation) Workbook.

3.13 Angular and Linear Measurement Analysis

The angular and linear measurements were recorded in an Excel© (Microsoft Corporation) spreadsheet and accuracy was recorded within 2° and 2mm respectively.

3.14 Data Analysis

All the data were recorded in an Excel© (Microsoft Corporation) Workbook. Statistical analysis was carried out using R© Version 4.1.0 (R Foundation for Statistical Computing Vienna, Austria, 2021). The statistical analysis included:

- Descriptive statistics were used to describe proportions, means and standard deviations
- Box Plots were constructed to illustrate the range of point-to-point error of each landmark and the range of differences between the x- and y-axis measurements for each landmark
- Bland-Altman Plots were constructed to illustrate the level of agreement for landmark identification between hand tracing and computerised cephalometric analysis for x- and y-axis coordinates of landmarks
- Intraclass correlation coefficients (ICC) and 95% confidence intervals were calculated to quantify inter-rater reliability. The mixed factorial design ICC was chosen as the rater(s) in this study were fixed.

Chapter 4 Results

4.1 Inter-Examiner Reliability

The results of the inter-examiner reliability were examined firstly by comparing the data for the three examiners together and then by comparing the data between each examiner pair (Examiner 1 versus Examiner 2, Examiner 1 versus Examiner 3, and Examiner 2 versus Examiner 3).

4.1.1 Three Examiners

Intraclass correlation coefficients (ICCs) were calculated for angular and linear measurements, x-axis landmark measurements, and y-axis landmark measurements for the three examiners. ICC values between 0.5 and 0.7 indicate moderate inter-rater agreement, values between 0.75 and 0.9 indicate good inter-rater agreement and values greater than 0.9 indicate excellent inter-rater agreement. For the inter-examiner angular and linear measurements, the ICC value for lower incisor to A-pogonion was 0.7 indicating moderate agreement, while the ICC values for the rest of the measurements ranged from 0.77-0.98, indicating good to excellent agreement. All the ICC values were statistically significant. The ICCs, 95% confidence intervals and p-values for this group are outlined in table 4.1.

The inter-examiner ICCs for x-axis measurements for both porion and orbitale were 0.86, indicating very good agreement and for the rest of the x-axis measurements are between 0.96 and 1.00, indicating excellent agreement between the three examiners. All the ICC values were statistically significant. The results are outlined in detail in table 4.2.

The inter examiner ICC for y-axis measurements for porion was 0.74 indicating moderate inter-rater agreement. The ICC values for the rest of the y-axis measurements were between 0.86 and 1.00, indicating good to excellent inter-examiner agreement. All the ICC values were statistically significant. The results are outlined in detail in table 4.3.

Table 4.1 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by Examiner 1, Examiner 2, and Examiner 3

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.92	0.81-0.98	<0.001
SNB	0.94	0.82-0.98	<0.001
ANB	0.78	0.52-0.93	<0.001
UI Mx	0.77	0.50-0.93	<0.001
LI Md	0.88	0.72-0.96	<0.001
II Angle	0.90	0.75-0.97	<0.001
MMPA	0.92	0.80-0.98	<0.001
Overjet	0.95	0.87-0.99	<0.001
LI-APog	0.70	0.39-0.90	<0.001
Downs			
Facial Angle	0.87	0.69-0.96	<0.001
Angle of Convexity	0.88	0.72-0.96	<0.001
FMPA	0.86	0.67-0.96	<0.001
Y-Axis	0.88	0.71-0.96	<0.001
LI-OcP	0.85	0.65-0.95	<0.001
UI-APog	0.98	0.94-0.99	<0.001

Table 4.2 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by Examiner 1, Examiner 2, and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.96	0.90-0.99	<0.001
N	0.99	0-98-1.00	<0.001
A	0.99	0.97-1.00	<0.001
B	1.00	1.00-1.00	<0.001
ANS	0.99	0.97-1.00	<0.001
PNS	0.96	0.88-0.99	<0.001
UIA	0.98	0.94-0.99	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.96	0.89-0.99	<0.001
LIT	1.00	0.99-1.00	<0.001
Pog	1.00	1.00-1.00	<0.001
Me	1.00	1.00-1.00	<0.001
Go	0.97	0.93-0.99	<0.001
Po	0.86	0.66-0.96	<0.001
Or	0.86	0.66-0.96	<0.001
Gn	1.00	0.99-1.00	<0.001

Table 4.3 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by Examiner 1, Examiner 2, and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.99	0.98-1.00	<0.001
N	0.96	0.90-0.99	<0.001
A	0.96	0.90-0.99	<0.001
B	0.96	0.89-0.99	<0.001
ANS	1.00	0.99-1.00	<0.001
PNS	0.86	0.67-0.96	<0.001
UIA	0.95	0.88-0.99	<0.001
UIT	0.98	0.95-0.99	<0.001
LIA	0.92	0.80-0.98	<0.001
LIT	0.99	0.97-1.00	<0.001
Pog	0.97	0.93-0.99	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.80	0.56-0.94	<0.001
Po	0.74	0.46-0.97	<0.001
Or	0.90	0.76-0.97	<0.001
Gn	0.99	0.97-1.00	<0.001

4.1.2 Examiner Pairs

Intraclass correlation coefficients (ICCs) were calculated for angular and linear measurements, x-axis measurements, and y-axis measurements for the three pairs of examiners: Examiner 1 and Examiner 2, Examiner 1 and Examiner 3, and Examiner 2 and Examiner 3.

4.1.3 Angular and Linear Measurements

The ICC values for angular and linear measurements recorded by Examiner 1 and Examiner 2 are shown in Table 4.4. The ICC values for these measurements ranged from 0.78 to 0.98 and were statistically significant, indicating good to excellent levels of agreement.

The ICC values for angular and linear measurements recorded by Examiner 1 and Examiner 3 are shown in Table 4.5. The ICC values for these measurements ranged from 0.75 to 0.97 and were statistically significant, indicating good to excellent levels of agreement.

The ICC values for angular and linear measurements recorded by Examiner 2 and Examiner 3 are shown in Table 4.6. The ICC for ANB is 0.72 and for lower incisor to A-pogonion was 0.74, indicating moderate inter-rater agreement between Examiner 2 and Examiner 3. The ICC values for the rest of these measurements ranged from 0.84 to 0.98, indicating good to excellent levels of agreement. All the ICCs were statistically significant.

Table 4.4 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by Examiner 1 and Examiner 2

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.94	0.78-0.98	<0.001
SNB	0.97	0.90-0.99	<0.001
ANB	0.82	0.45-0.95	<0.001
UI Mx	0.89	0.64-0.97	<0.001
LI Md	0.92	0.73-0.98	<0.001
II Angle	0.86	0.55-0.96	<0.001
MMPA	0.91	0.69-0.97	<0.001
Overjet	0.95	0.82-0.99	<0.001
LI-APog	0.75	0.30-0.93	<0.05
Downs			
Facial Angle	0.79	0.37-0.94	<0.01
Angle of Convexity	0.87	0.58-0.96	<0.01
FMPA	0.80	0.42-0.94	<0.01
Y-Axis	0.83	0.49-0.95	<0.01
LI-OcP	0.77	0.33-0.93	<0.01
UI-APog	0.97	0.90-0.99	<0.001

Table 4.5 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by Examiner 1 and Examiner 3

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.88	0.61-0.97	<0.001
SNB	0.91	0.70-0.98	<0.001
ANB	0.82	0.45-0.95	<0.01
UI Mx	0.86	0.55-0.96	<0.01
LI Md	0.92	0.73-0.98	<0.001
II Angle	0.89	0.65-0.97	<0.001
MMPA	0.90	0.68-0.97	<0.001
Overjet	0.94	0.78-0.98	<0.001
LI-APog	0.88	0.62-0.97	<0.001
Downs			
Facial Angle	0.97	0.91-0.99	<0.001
Angle of Convexity	0.87	0.57-0.96	<0.01
FMPA	0.93	0.76-0.98	<0.001
Y-Axis	0.95	0.81-0.99	<0.001
LI-OcP	0.78	0.37-0.94	<0.01
UI-APog	0.98	0.94-1.00	<0.001

Table 4.6 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by Examiner 2 and Examiner 3

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.95	0.82-0.99	<0.001
SNB	0.95	0.82-0.99	<0.001
ANB	0.72	0.24-0.92	<0.05
UI Mx	0.88	0.61-0.97	<0.001
LI Md	0.98	0.93-0.99	<0.001
II Angle	0.94	0.78-0.98	<0.001
MMPA	0.96	0.86-0.99	<0.001
Overjet	0.97	0.90-0.99	<0.001
LI-APog	0.74	0.28-0.92	<0.05
Downs			
Facial Angle	0.84	0.50-0.95	<0.05
Angle of Convexity	0.92	0.72-0.98	<0.001
FMPA	0.86	0.54-0.96	<0.01
Y-Axis	0.84	0.51-0.96	<0.01
LI-OcP	0.98	0.92-0.99	<0.001
UI-APog	0.98	0.91-0.99	<0.001

4.1.4 X- and Y-Axis Measurements Examiner 1 vs. Examiner 2

The ICC values for x-axis landmark measurements recorded by Examiner 1 and Examiner 2 are shown in Table 4.7. They range from 0.84 to 1.00 and were all statistically significant, indicating very good to excellent inter-rater reliability. The remaining ICC values for this group range from 0.77 to 0.97 and were statistically significant.

Table 4.7 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by Examiner 1 and Examiner 2

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.95	0.83-0.99	<0.001
N	1.00	0.98-1.00	<0.001
A	1.00	0.98-1.00	<0.001
B	1.00	0.99-1.00	<0.001
ANS	0.99	0.98-1.00	<0.001
PNS	0.92	0.71-0.98	<0.001
UIA	0.98	0.94-1.00	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.98	0.97-1.00	<0.001
LIT	1.00	0.99-1.00	<0.001
Pog	1.00	1.00-1.00	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.96	0.85-0.99	<0.001
Po	0.84	0.49-0.95	<0.01
Or	0.84	0.49-0.95	<0.01
Gn	1.00	0.99-1.00	<0.001

The ICC values for y-axis landmark measurements recorded by Examiner 1 and Examiner 2 are shown in Table 4.8. The ICC value for gonion was 0.67 and was statistically significant, indicating moderate inter-rater reliability. The rest of the ICC values ranged from 0.77 to 1.00 and were all statistically significant, indicating good to excellent inter-rater reliability.

Table 4.8 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by Examiner 1 and Examiner 2

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	1.00	0.99-1.00	<0.001
N	0.97	0.90-0.99	<0.001
A	0.95	0.84-0.99	<0.001
B	0.95	0.84-0.99	<0.001
ANS	1.00	0.99-1.00	<0.001
PNS	0.83	0.48-0.95	<0.01
UIA	0.92	0.74-0.98	<0.001
UIT	0.98	0.92-0.99	<0.001
LIA	0.90	0.68-0.97	<0.001
LIT	0.99	0.96-1.00	<0.001
Pog	0.96	0.86-0.99	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.67	0.15-0.90	<0.05
Po	0.77	0.34-0.93	<0.01
Or	0.88	0.62-0.97	<0.001
Gn	1.00	1.00-1.00	<0.001

4.1.5 X- and Y-Axis Measurements Examiner 1 vs. Examiner 3

The ICC values for x-axis landmark measurements recorded by Examiner 1 and Examiner 3 are shown in Table 4.9. They ranged from 0.86 to 1.00 and were all statistically significant, indicating very good to excellent inter-rater reliability.

Table 4.9 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by Examiner 1 and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.98	0.93-0.99	<0.001
N	0.99	0.97-1.00	<0.001
A	0.99	0.95-1.00	<0.001
B	1.00	1.00-1.00	<0.001
ANS	0.99	0.95-1.00	<0.001
PNS	0.97	0.90-0.99	<0.001
UIA	0.98	0.94-1.00	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.95	0.82-0.99	<0.001
LIT	1.00	0.99-1.00	<0.001
Pog	1.00	1.00-1.00	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.98	0.93-1.00	<0.001
Po	0.87	0.57-0.96	<0.01
Or	0.86	0.57-0.96	<0.01
Gn	1.00	0.99-1.00	<0.001

The ICC values for y-axis landmark measurements recorded by Examiner 1 and Examiner 3 are shown in Table 4.10. They ranged from 0.77 to 1.00 and were all statistically significant, indicating good to excellent inter-rater reliability.

Table 4.10 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by Examiner 1 and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.99	0.95-1.00	<0.001
N	0.94	0.80-0.98	<0.001
A	0.96	0.84-0.99	<0.001
B	0.98	0.94-1.00	<0.001
ANS	1.00	0.98-1.00	<0.001
PNS	0.84	0.51-0.96	<0.01
UIA	0.97	0.88-0.99	<0.001
UIT	0.99	0.96-1.00	<0.001
LIA	0.90	0.67-0.97	<0.001
LIT	0.99	0.95-1.00	<0.001
Pog	0.99	0.96-1.00	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.93	0.75-0.98	<0.001
Po	0.77	0.33-0.93	<0.01
Or	0.88	0.61-0.97	<0.001
Gn	0.99	0.96-1.00	<0.001

4.1.6 X- and Y-Axis Measurements Examiner 2 vs. Examiner 3

The ICC values for x-axis landmark measurements recorded by Examiner 2 and Examiner 3 are shown in Table 4.9. They ranged from 0.87 to 1.00 and were all statistically significant, indicating very good to excellent inter-rater reliability.

Table 4.11 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by Examiner 2 and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.95	0.83-0.99	<0.001
N	0.99	0.96-1.00	<0.001
A	0.99	0.95-1.00	<0.001
B	1.00	0.99-1.00	<0.001
ANS	0.98	0.94-1.00	<0.001
PNS	0.98	0.91-0.99	<0.001
UIA	0.97	0.89-0.99	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.99	0.98-1.00	<0.001
LIT	0.99	0.98-1.00	<0.001
Pog	1.00	1.00-1.00	<0.001
Me	1.00	1.00-1.00	<0.001
Go	0.98	0.93-1.00	<0.001
Po	0.87	0.57-0.96	<0.01
Or	0.88	0.61-0.97	<0.001
Gn	1.00	0.99-1.00	<0.001

The ICC values for y-axis landmark measurements recorded by Examiner 2 and Examiner 3 are shown in Table 4.9. They ranged from 0.85 to 1.00 and were all statistically significant, indicating very good to excellent inter-rater reliability.

Table 4.12 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by Examiner 2 and Examiner 3

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.99	0.95-1.00	<0.001
N	0.97	0.88-0.99	<0.001
A	0.98	0.93-0.99	<0.001
B	0.94	0.79-0.98	<0.001
ANS	1.00	0.99-1.00	<0.001
PNS	0.97	0.89-0.99	<0.001
UIA	0.97	0.89-0.99	<0.001
UIT	0.98	0.91-0.99	<0.001
LIA	0.97	0.89-0.99	<0.001
LIT	0.99	0.97-1.00	<0.001
Pog	0.98	0.91-0.99	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.85	0.54-0.96	<0.001
Po	0.87	0.57-0.96	<0.01
Or	0.98	0.92-0.99	<0.001
Gn	0.99	0.97-1.00	<0.001

4.1.7 Point-to-Point Measurements Between Examiner Pairs

Comparing each examiner pair, 84.38% of the landmarks were identified within 2mm of each other for Examiner 1 and Examiner 2, 87.5% of the landmarks were identified within 2mm of each other for Examiner 1 and Examiner 3 and 84.38% of the landmarks were identified within 2mm of each other for Examiner 2 and Examiner 3. The two least accurately identified landmarks across all three pairs of examiners are A-point and porion.

Table 4.13 Point to Point (PtoP) measurement proportions less than 2mm for each landmark between each pair of examiners: Examiner 1 and Examiner 2 (E1 vs E2), Examiner 1 and Examiner 3 (E1 vs E3), and Examiner 2 and Examiner 3 (E2 vs E3)

Landmark	PtoP<2mm E1 vs E2	PtoP<2mm E1 vs E3	PtoP<2mm E2 vs E3
S	100.00	100.00	100.00
N	87.50	87.50	75.00
A	62.50	62.50	37.50
B	100.00	100.00	87.50
ANS	100.00	100.00	100.00
PNS	87.50	62.50	100.00
UIA	75.00	75.00	62.50
UIT	100.00	100.00	100.00
LIA	50.00	62.50	100.00
LIT	100.00	100.00	100.00
Pog	75.00	100.00	100.00
Me	100.00	100.00	100.00
Go	100.00	87.50	62.50
Po	50.00	87.50	37.50
Or	62.50	75.00	87.50
Gn	100.00	100.00	100.00
AVERAGE	**84.38**	**87.50**	**84.38**

4.2 Intra-Examiner Reliability

The results of the intra-examiner reliability were examined by comparing the measurements recorded by a single examiner on two separate occasions. Intraclass correlation coefficients (ICCs) were calculated for angular and linear measurements, x-axis landmark measurements, and y-axis landmark measurements.

4.2.1 Angular and Linear Measurements

The ICC values for angular and linear measurements recorded by a single examiner at two different time points are shown in Table 4.14. The ICC values for these measurements ranged from 0.87 to 0.99 and were statistically significant, indicating good to excellent levels of agreement.

4.2.2 X- and Y-Axis Measurements

The ICC values for x-axis landmark measurements recorded by a single examiner at two different time points are shown in Table 4.15. They ranged from 0.94 to 1.00 and were all statistically significant, indicating excellent levels of agreement.

The ICC values for y-axis landmark measurements recorded by a single examiner at two different time points are shown in Table 4.16. They ranged from 0.87 to 1.00 and were all statistically significant, indicating good to excellent levels of agreement.

Table 4.14 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by a single examiner at two different time points

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.95	0.86-0.98	<0.001
SNB	0.99	0.96-0.99	<0.001
ANB	0.92	0.80-0.97	<0.001
UI Mx	0.98	0.94-0.99	<0.001
LI Md	0.91	0.77-0.97	<0.001
II Angle	0.99	0.97-1.00	<0.001
MMPA	0.98	0.94-0.99	<0.001
Overjet	0.99	0.98-1.00	<0.001
LI-APog	0.94	0.85-0.98	<0.001
Downs			
Facial Angle	0.87	0.69-0.95	<0.001
Angle of Convexity	0.94	0.84-0.98	<0.001
FMPA	0.91	0.79-0.97	<0.001
Y-Axis	0.90	0.76-0.96	<0.001
LI-OcP	0.93	0.82-0.97	<0.001
UI-APog	0.97	0.93-0.99	<0.001

Table 4.15 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by a single examiner at two different time points

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	0.99	0.98-1.00	<0.001
N	0.99	0.97-1.00	<0.001
A	1.00	0.99-1.00	<0.001
B	1.00	1.00-1.00	<0.001
ANS	0.98	0.95-0.99	<0.001
PNS	0.97	0.93-0.99	<0.001
UIA	0.99	0.99-1.00	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.95	0.88-0.98	<0.001
LIT	1.00	1.00-1.00	<0.001
Pog	1.00	1.00-1.00	<0.001
Me	1.00	0.99-1.00	<0.001
Go	0.98	0.95-0.99	<0.001
Po	0.96	0.88-0.98	<0.001
Or	0.94	0.85-0.98	<0.001
Gn	0.99	0.98-1.00	<0.001

Table 4.16 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by a single examiner at two different time points

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
S	1.00	0.99-1.00	<0.001
N	1.00	1.00-1.00	<0.001
A	1.00	0.99-1.00	<0.001
B	0.98	0.96-0.99	<0.001
ANS	1.00	0.99-1.00	<0.001
PNS	0.99	0.99-1.00	<0.001
UIA	0.99	0.97-1.00	<0.001
UIT	1.00	1.00-1.00	<0.001
LIA	0.99	0.96-0.99	<0.001
LIT	1.00	1.00-1.00	<0.001
Pog	0.98	0.95-0.99	<0.001
Me	1.00	1.00-1.00	<0.001
Go	0.95	0.86-0.98	<0.001
Po	0.87	0.69-0.95	<0.001
Or	0.99	0.97-1.00	<0.001
Gn	1.00	1.00-1.00	<0.001

4.2.3 Point-to-Point Measurements

Comparing the measurements recorded by a single examiner on two separate occasions, 91.83% of the landmarks were identified within 2mm of each other. The individual proportions for each landmark are shown in Table 4.17 and indicate a high level of agreement.

Table 4.17 Point-to-point measurement proportions less than 2mm for a single examiner at two different time points

Landmark	Point-to-Point <2mm (%)
S	100.00
N	92.31
A	100.00
B	84.62
ANS	84.62
PNS	92.31
UIA	92.31
UIT	100.00
LIA	92.31
LIT	100.00
Pog	84.62
Me	92.31
Go	84.62
Po	92.31
Or	76.92
Gn	100.00
AVERAGE	**91.83**

4.3 Comparison of hand tracing and computerised cephalometric analysis

The successful detection rates (SDR) for 1.0mm, 2.0mm, 2.5mm, 3.0mm and 4.0mm are described in Table 4.18. The SDR values are shown for each individual landmark along with the average SDRs for each precision range. The average SDR value for each precision range is also illustrated graphically in Figure 4.1. The most relevant result is the average SDR for the 2.0mm precision range, which was 82.08%. This finding means that 82.08% of the automatically located landmarks were identified within 2mm of the hand traced landmark, i.e., the gold standard. The mean radial error (MRE) for each landmark is shown in Table 4.18 and the average MRE was 1.25 ± 0.91mm.

Figure 4.1 Successful Detection Rates (SDR) for 1.0mm, 2.0mm, 2.5mm, 3.0mm and 3.5mm precision ranges

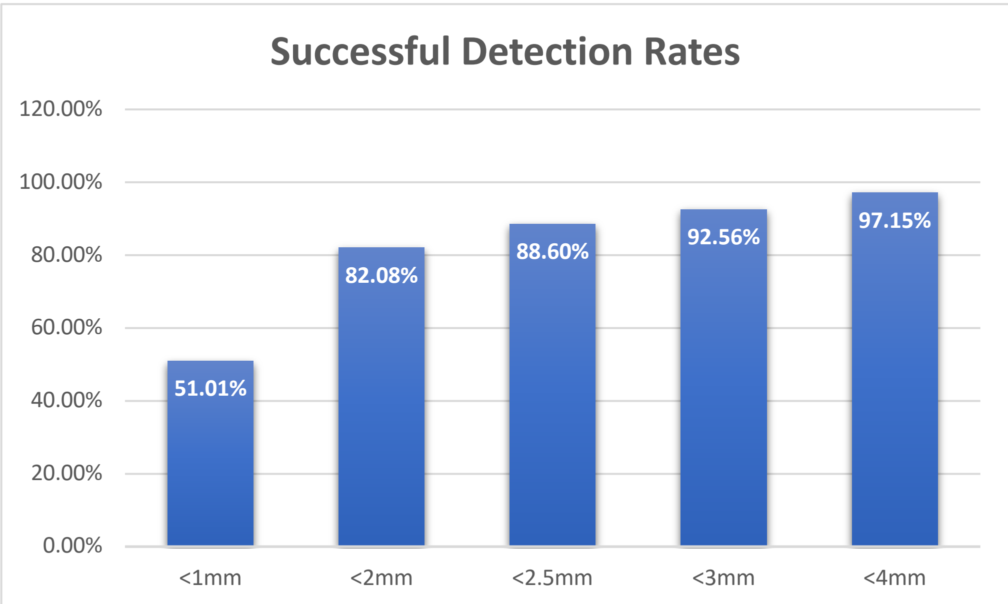
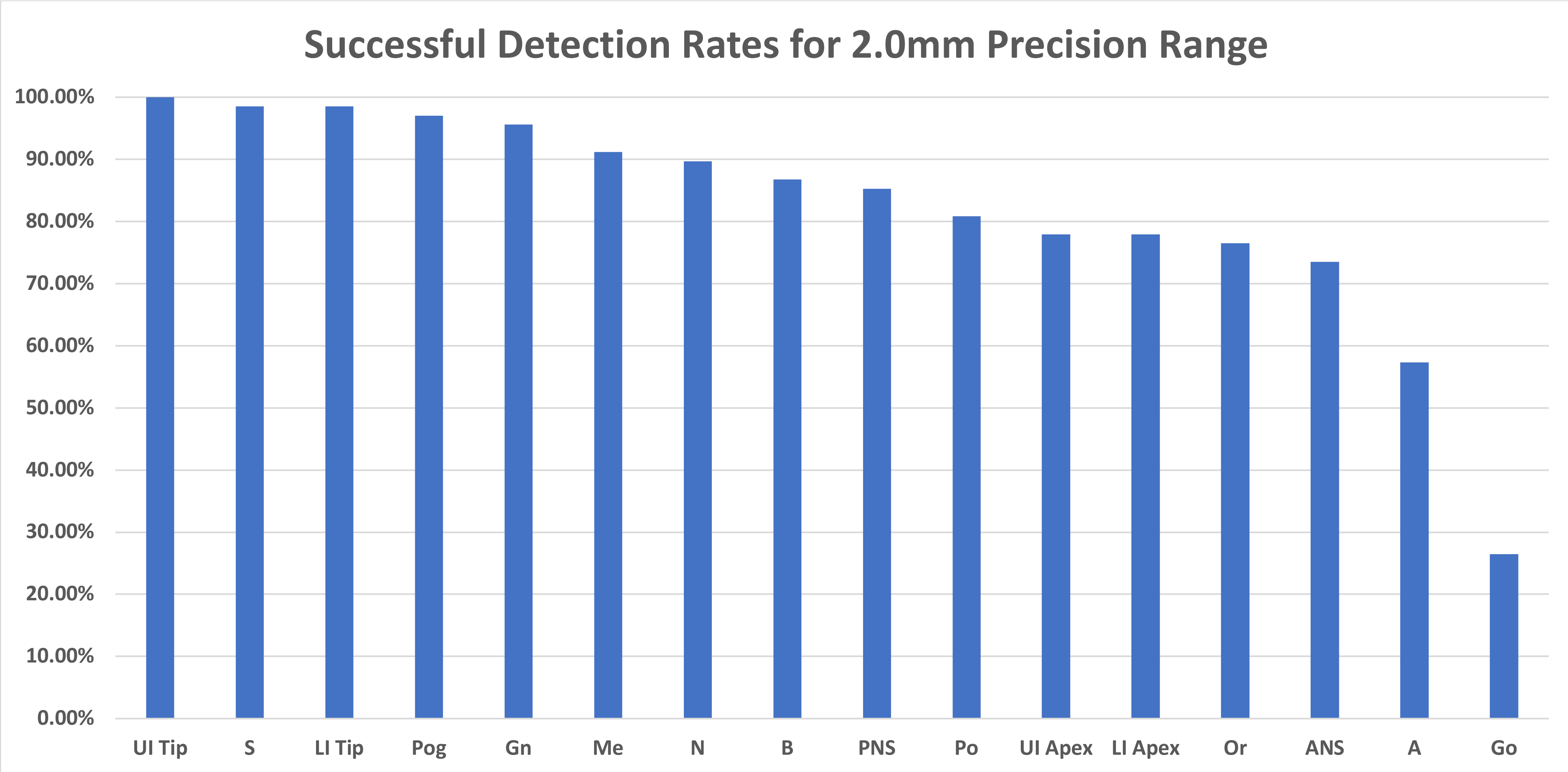


Table 4.18 Point-to-point errors for each individual landmark and successful detection rates (SDR) for 1.0mm, 2.0mm, 2.5mm, 3.0mm and 4.0mm precision ranges

Landmark	MRE±SD(mm)	SDR 1.0mm	SDR 2.0mm	SDR 2.5mm	SDR 3.0mm	SDR 4.0mm
S	0.79 ± 0.57	67.65	98.53	98.53	98.53	100.00
N	0.98 ± 1.17	69.12	89.71	92.65	92.65	97.00
A	1.77 ± 1.06	23.53	57.35	73.53	86.76	97.00
B	1.02 ± 0.82	54.41	86.76	95.59	98.53	99.00
ANS	1.42 ± 1.05	41.18	73.53	89.71	95.59	99.00
PNS	1.26 ± 0.82	41.18	85.29	92.65	95.59	100.00
UIA	1.33 ± 0.99	48.53	77.94	85.29	92.65	99.00
UIT	0.77 ± 0.38	69.12	100.00	100.00	100.00	100.00
LIA	1.38 ± 1.03	39.71	77.94	86.76	94.12	99.00
LIT	0.71 ± 0.48	67.65	98.53	100.00	100.00	100.00
Pog	0.71 ± 0.58	75.00	97.06	97.06	100.00	100.00
Me	1.10 ± 0.71	45.59	91.18	95.59	97.06	100.00
Go	3.24 ± 1.86	5.88	26.47	42.65	52.94	74.00
Po	1.42 ± 1.21	42.65	80.88	86.76	89.71	97.00
Or	1.42 ± 1.28	51.47	76.47	82.35	86.76	96.00
Gn	0.69 ± 0.59	73.53	95.59	98.53	100.00	100.00
AVERAGE	**1.25 ± 0.91**	**51.01**	**82.08**	**88.60**	**92.56**	**97.15**

The SDR for each landmark in the 2.0mm precision range is illustrated graphically in Figure 4.2. Upper incisor tip was the most frequently successfully identified landmark with an SDR of 100% and gonion was the least frequently successfully identified landmark with an SDR of 26.47%. This is followed by A-Point, which had an SDR of 57.35% in the 2.0mm precision range. The landmarks with lower SDR values in the 2.0mm precision range substantially increased their SDR values in the 2.5mm precision range with gonion increasing from 26.47% at 2.0mm to 52.65% at 2.5mm and A-point increasing from 57.35% at 2.0mm to 73.53% at 2.5mm.

Figure 4.2 Successful Detection Rate (SDR) for 2.0mm precision range for each landmark



Out of the 16 landmarks identified in this study, nine were successfully identified within a 2.0mm precision range with an accuracy of greater than 85%: upper incisor tip (100%), sella (98.53%), lower incisor tip (98.53%), pogonion (97.06%), gnathion (95.59%), menton (91.18%), nasion (89.71%), B-point (86.76%) and posterior nasal spine (85.29%).

Five of the 16 landmarks were successfully identified within a 2.0mm precision range with an accuracy ranging from 73-81%: porion (80.88%), upper incisor apex (77.94%), lower incisor apex (77.94%) orbitale (76.47%) and anterior nasal spine (73.53%).

4.4 Angular and Linear Measurements

The accuracy of the angular and linear measurements was recorded within 2° or 2mm for angular and linear measurements respectively. These results are shown in Table 4.19. The most accurate measurement was overjet, with 100% of the automated measurements recorded within 2mm of the hand traced measurements. The least accurate measurement was lower incisor to occlusal plane angle, with only 50% of the automated measurements recorded within 2°. This was closely followed by upper incisor to maxillary plane angle with an accuracy of 54.41%, lower incisor to mandibular plane angle with an accuracy of 55.88% and interincisal angle with an accuracy of 57.35% of the automated measurements recorded within 2° of the hand traced measurements. The rest of the angular and linear measurements produced accuracies of greater than 80%. The average success rate for all measurements was 80.98%.

Comparing the Eastman and the Downs analysis in terms of accuracy, the Eastman analysis was more accurate with an average success rate of 81.05%. The Downs analysis had an average success rate of 74.82%. These results are shown in detail in table 4.20.

Table 4.19 Accuracy of Eastman analysis and Downs analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared to traditional hand tracing

Analysis	Angle	Accuracy (%)	Analysis	Angle	Accuracy (%)
Eastman	SNA	92.65%	**Downs**	Facial Angle	88.24%
	SNB	92.65%		Angle of Convexity	82.35%
	ANB	98.53%		FMPA	82.35%
	UI Mx	54.41%		Y-Axis	89.71%
	LI Md	55.88%		LI/OcP	50.00%
	II Angle	57.35%		UI/A-Pog	92.65%
	MMPA	86.76%		LI Md	55.88%
	Overjet	100.00%		II Angle	57.35%
	LI/A-Pog	91.18%			
AVERAGE		**81.05%**	**AVERAGE**		**74.82%**

Table 4.20 Proportion of angular and linear measurement recorded within 2°/2mm

Analysis	Error <2°/2mm (%)
Eastman	
SNA	92.65
SNB	92.65
ANB	98.53
UI Mx	54.41
LI Md	55.88
II Angle	57.35
MMPA	88.76
Overjet	100.00
LI-APog	91.18
Downs	
Facial Angle	88.24
Angle of Convexity	82.35
FMPA	82.35
Y-Axis	89.71
LI-OcP	50.00
UI-APog	92.65
AVERAGE	**80.98**

4.5.1 Intraclass Correlation Coefficients

Intraclass correlation coefficients (ICCs) were calculated for angular and linear measurements, x-axis landmark measurements, and y-axis landmark measurements.

The ICC values for angular and linear measurements recorded by the principal examiner and Planmeca Romexis® are shown in Table 4.21. The ICC value for lower incisor to A-Pog is 0.73 and is statistically significant, indicating moderate inter-rater reliability. The rest of the ICC values range from 0.86 to 0.97 and are all statistically significant, indicating good to excellent inter-rater reliability.

Table 4.21 Intraclass Correlation Coefficients (ICCs) for angular and linear measurements recorded by hand tracing and Planmeca Romexis®

Eastman	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
SNA	0.92	0.88-0.94	<0.001
SNB	0.94	0.92-0.96	<0.001
ANB	0.96	0.95-0.98	<0.001
UI Mx	0.94	0.91-0.96	<0.001
LI Md	0.92	0.89-0.95	<0.001
II Angle	0.95	0.93-0.97	<0.001
MMPA	0.97	0.96-0.98	<0.001
Overjet	0.97	0.96-0.98	<0.001
LI-APog	0.73	0.62-0.81	<0.001
Downs			
Facial Angle	0.86	0.80-0.91	<0.001
Angle of Convexity	0.97	0.95-0.98	<0.001
FMPA	0.94	0.91-0.96	<0.001
Y-Axis	0.93	0.90-0.95	<0.001
LI-OcP	0.90	0.85-0.93	<0.001
UI-APog	0.97	0.95-0.98	<0.001

4.5.2 X-Axis Measurements

X- and y-axis measurements were recorded for the nine least accurately identified landmarks: A-point, B-point, anterior nasal spine, posterior nasal spine, upper incisor apex, lower incisor apex, gonion, porion and orbitale. The ICC values for x-axis landmark measurements recorded by the principal investigator and Planmeca Romexis® are shown in Table 4.22. They ranged from 0.90 to 1.00 and are all statistically significant, indicating very good to excellent inter-rater reliability.

Table 4.22 Intraclass Correlation Coefficients (ICCs) for x-axis landmark measurements recorded by hand tracing and Planmeca Romexis®

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
A	0.99	0.98-0.99	<0.001
B	1.00	0.99-1.00	<0.001
ANS	0.96	0.93-0.97	<0.001
PNS	0.93	0.90-0.95	<0.001
UIA	0.96	0.94-0.97	<0.001
LIA	0.98	0.96-0.98	<0.001
Go	0.88	0.82-0.92	<0.001
Po	0.96	0.94-0.97	<0.001
Or	0.90	0.86-0.93	<0.001

4.5.3 Y-Axis Measurements

The ICC values for y-axis landmark measurements recorded by the principal investigator and Planmeca Romexis® are shown in Table 4.22. They ranged from 0.94 to 0.99 and are all statistically significant, indicating excellent inter-rater reliability.

Table 4.23 Intraclass Correlation Coefficients (ICCs) for y-axis landmark measurements recorded by hand tracing and Planmeca Romexis®

Landmarks	Intraclass Correlation Coefficient	95% Confidence Interval	p-value
A	0.98	0.97-0.99	<0.001
B	0.99	0.98-0.99	<0.001
ANS	0.99	0.99-0.99	<0.001
PNS	0.98	0.98-0.99	<0.001
UIA	0.98	0.97-0.98	<0.001
LIA	0.98	0.98-0.99	<0.001
Go	0.95	0.93-0.97	<0.001
Po	0.94	0.91-0.96	<0.001
Or	0.96	0.94-0.97	<0.001

As the ICC values for the x- and y-axis measurements indicate a high level of agreement, but the landmarks were less frequently successfully identified, Bland-Altman plots were constructed to further illustrate the level of agreement. These are outlined below. For each plot, the x-axis represents the mean of the compared measurements, and the y-axis represents the difference between the compared measurements. The black horizontal line represents the mean difference between the two methods, or the bias. The red dashed lines represent the confidence interval of agreement. The greater the distance between the two red dashed lines, the greater variation between the landmark accuracy. Bland-Altman plots for the two least accurately identified landmarks, A-point and gonion, are included here in Figures 4.3-4.6. The remaining Bland-Altman plots are included in Appendix 8.

4.6 Bland-Altman Plots

Figure 4.3 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for A-Point

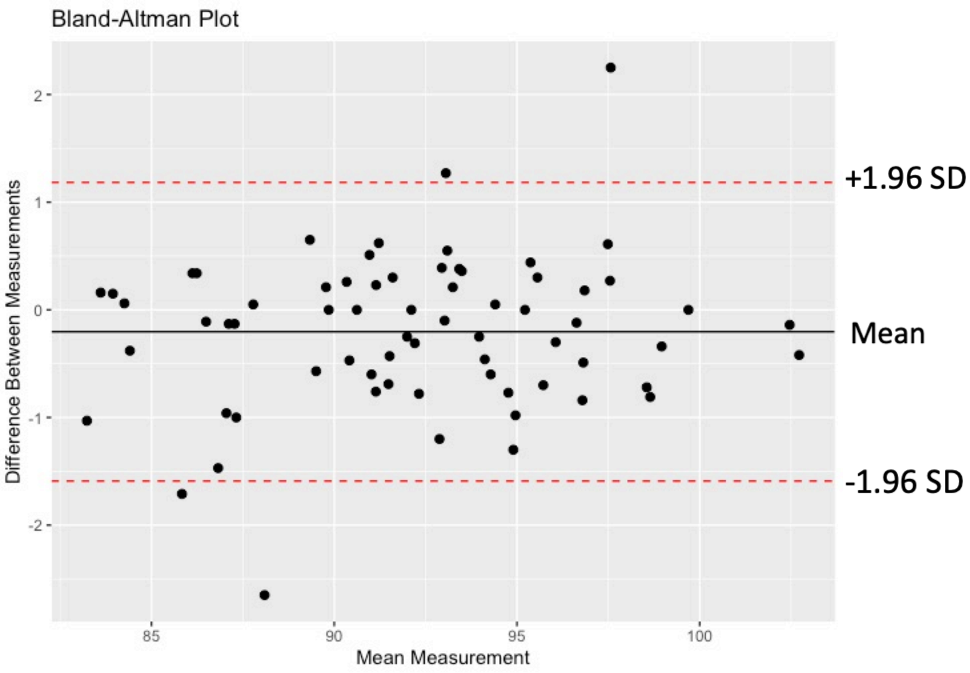
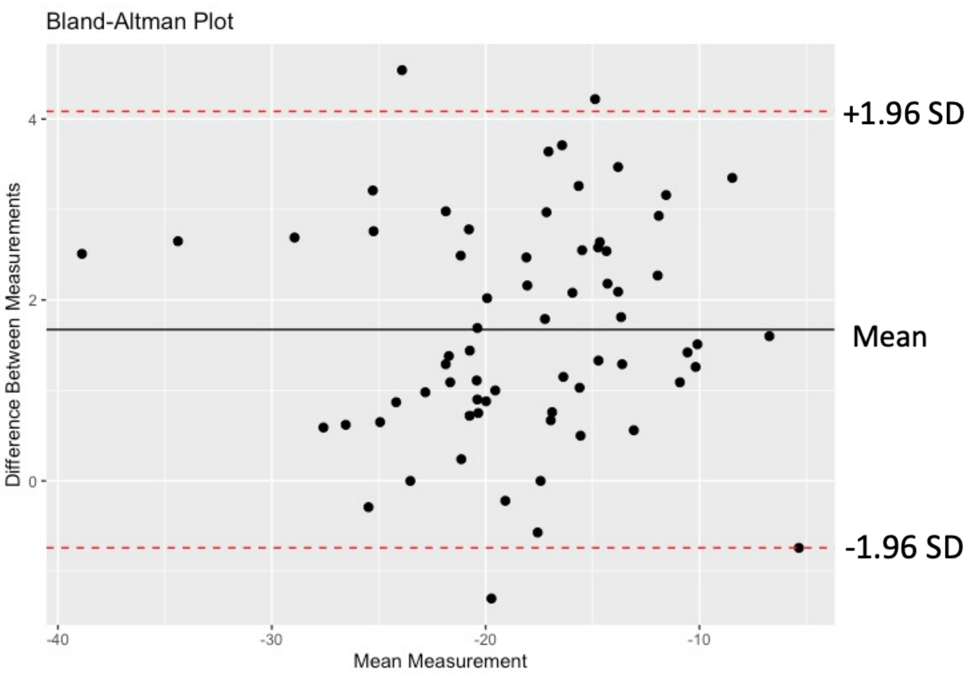


Figure 4.4 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for A-Point



Figures 4.3 and 4.4 show that the vertical discrepancy was greater than the horizontal discrepancy for the identification of A-Point by hand tracing and automated tracing.

Figure 4.5 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for gonion (Go)

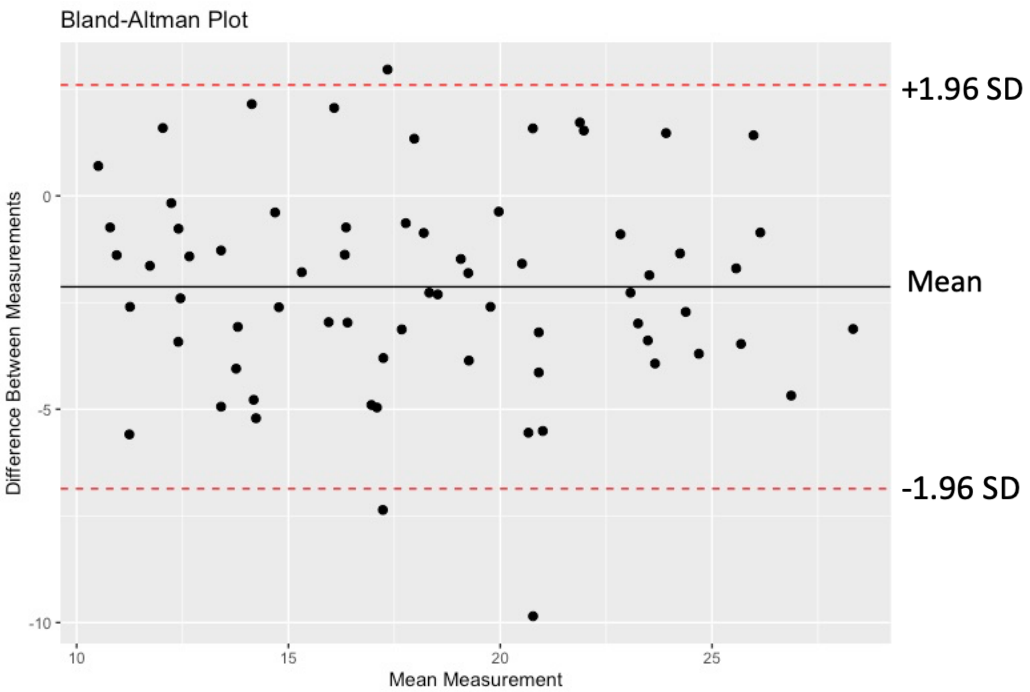
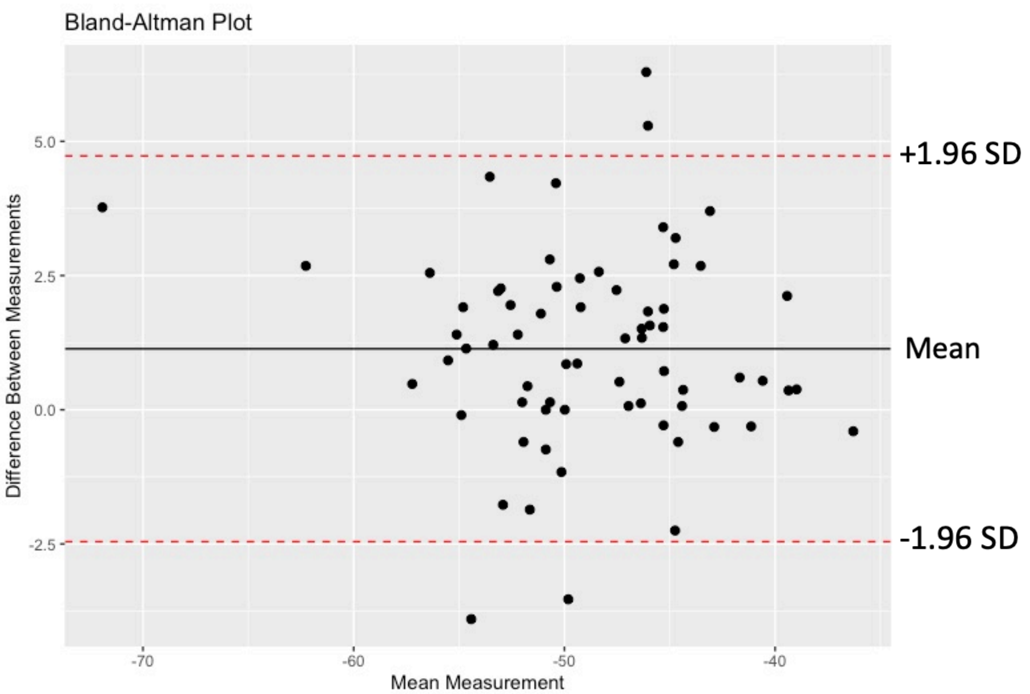


Figure 4.6 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for gonion (Go)



Figures 4.5 and 4.6 show that the horizontal and vertical discrepancies were both increased for the identification of gonion by hand tracing and automated tracing and that the vertical discrepancy was greater than the horizontal discrepancy.

4.7 Boxplots

Boxplots were constructed to illustrate the distribution of the data around the median and inter quartile range for each landmark and to visualise outliers. Figure 4.7 is the boxplot for the point-to-point error measurements between landmarks identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software and by traditional hand tracing. It shows that the greatest variation in point-to-point errors was recorded for gonion with outliers of up to 10mm. This is followed by variation in the distribution of the point-to-point errors for A-point.

Figure 4.7 Boxplot for the point-to-point error between landmarks identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software and by traditional hand tracing

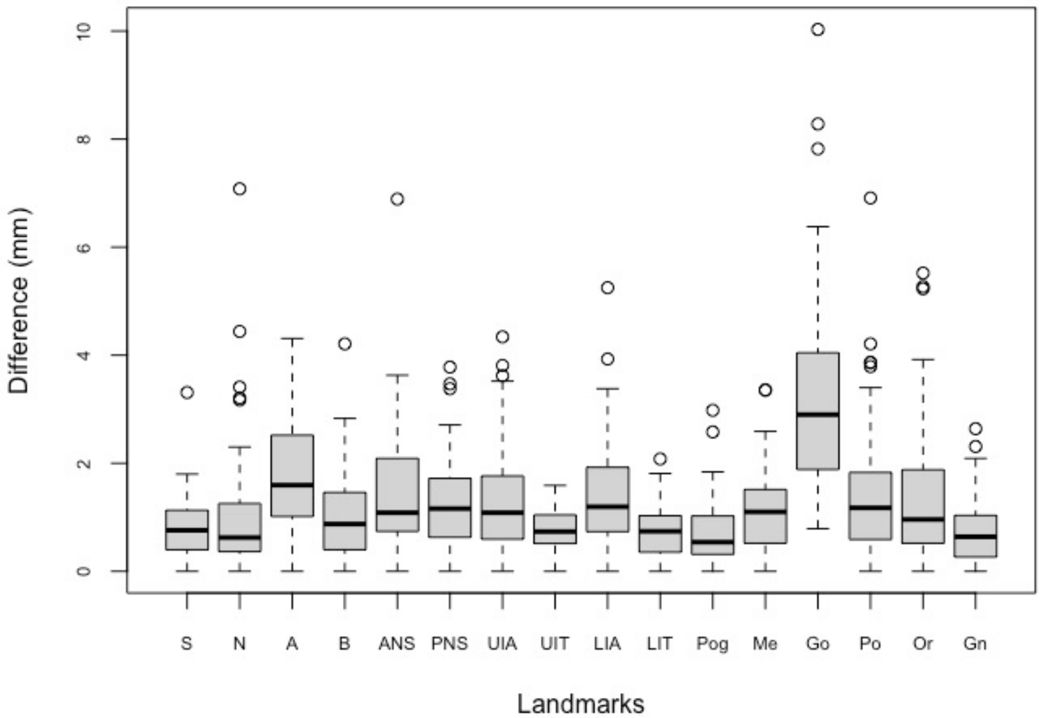


Figure 4.8 is the boxplot for the x-axis measurement differences for the nine landmarks that were least accurately identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared with traditional hand tracing. It shows that the greatest variation in the horizontal direction was recorded for gonion and that Planmeca Romexis® (Version 5.1.1.R) dental imaging software often located gonion too anteriorly compared to the landmarks located by the principal investigator. This is evident as most of the data for gonion are negative in value.

Figure 4.8 Boxplot for the x-axis measurement differences for the nine landmarks that were least accurately identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared with traditional hand tracing

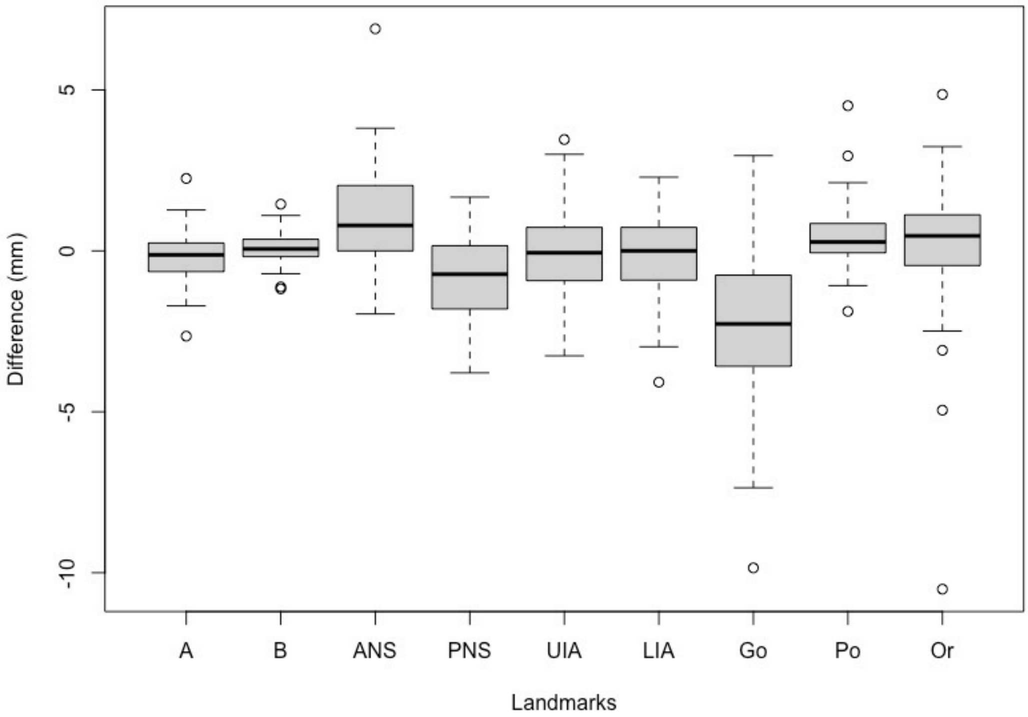
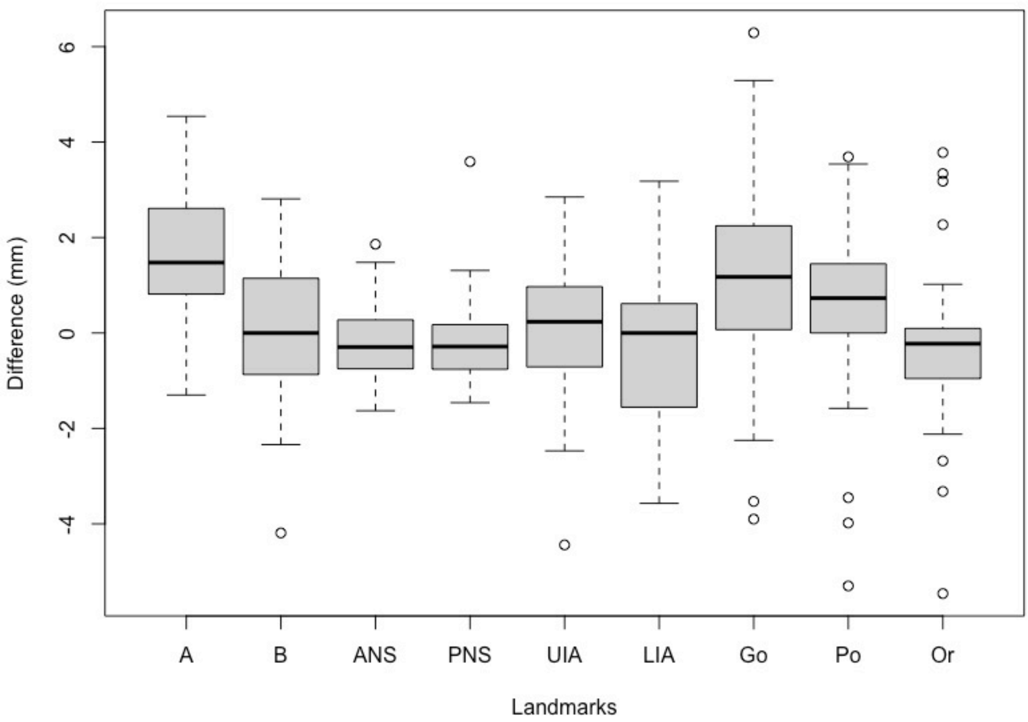


Figure 4.9 is the boxplot for the y-axis measurement differences for the 9 landmarks that were least accurately identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared with traditional hand tracing. It shows that several landmarks have a wide distribution of datapoints in the vertical direction: gonion, A-point, B-point, upper incisor apex, lower incisor apex and porion. For most of the landmarks, the direction of the vertical discrepancy is even. However, for A-point and gonion, the boxplot shows that Planmeca Romexis® (Version 5.1.1.R) dental imaging software often located these too inferiorly compared to the landmarks located by the principal investigator. This is evident as most of the data for gonion and A-point are positive in value.

Figure 4.9 Boxplot for the y-axis measurement differences for the nine landmarks that were least accurately identified automatically with Planmeca Romexis® (Version 5.1.1.R) dental imaging software compared with traditional hand tracing



Chapter 5 Discussion

Automated cephalometric analysis software programmes have been developed to reduce the time required to obtain an analysis and to reduce the risk of errors due to clinician's subjectivity, therefore improving the accuracy and reliability of the analysis.[5, 13] While these software programmes are significantly faster than traditional hand tracing or semi-automatic cephalometric software programmes in performing a cephalometric analysis, there is very limited evidence available on the accuracy of commercially available automated cephalometric analysis software programmes. However, there is a body of evidence relating to the accuracy of computer programming techniques that are capable of producing an automated cephalometric analysis. Recent published literature examining the efficacy of these computer programming techniques suggests that there have been significant improvements in their accuracy, and several are now capable of producing automated cephalometric analyses which are comparable with the gold standard, human expert measurement.

This aim of this study was to determine if automated computerised cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software is as accurate as traditional hand tracing in terms of landmark identification, angular and linear measurements. Inter- and intra-examiner reliability investigations were carried out prior to hand tracing the main study sample.

5.1 Overview of Key Findings

Good to excellent inter-examiner reliability was established for cephalometric measurements and for landmark identification relative to the x- and y-axes. The point-to-point measurements for all three examiner pairs revealed that 84.38-87.5% of landmarks were identified within 2mm of each other.

Very good to excellent intra-examiner reliability was established for cephalometric measurements and for landmark identification relative to the x- and y-axes. The point-to-point measurements revealed that 91.83% of landmarks recorded by the principal investigator on two separate occasions were identified within 2mm of each other.

The comparison of traditional hand tracing by the principal investigator and automated computerised cephalometric analysis by Planmeca Romexis® (Version 5.1.1.R) dental imaging software found that the average successful detection rate (SDR) of the automatically detected landmarks within a 2mm precision range of the hand traced landmarks was 82.08%. The mean radial error was 1.25 ± 0.91. In terms of the accuracy of the angular and linear measurements, 80.98% of the automated measurements were recorded within 2° or 2mm of the hand traced measurements.

Good to excellent agreement was established between the principal investigator and Planmeca Romexis® dental imaging software for cephalometric measurements and for landmark identification relative to the x- and y-axes.

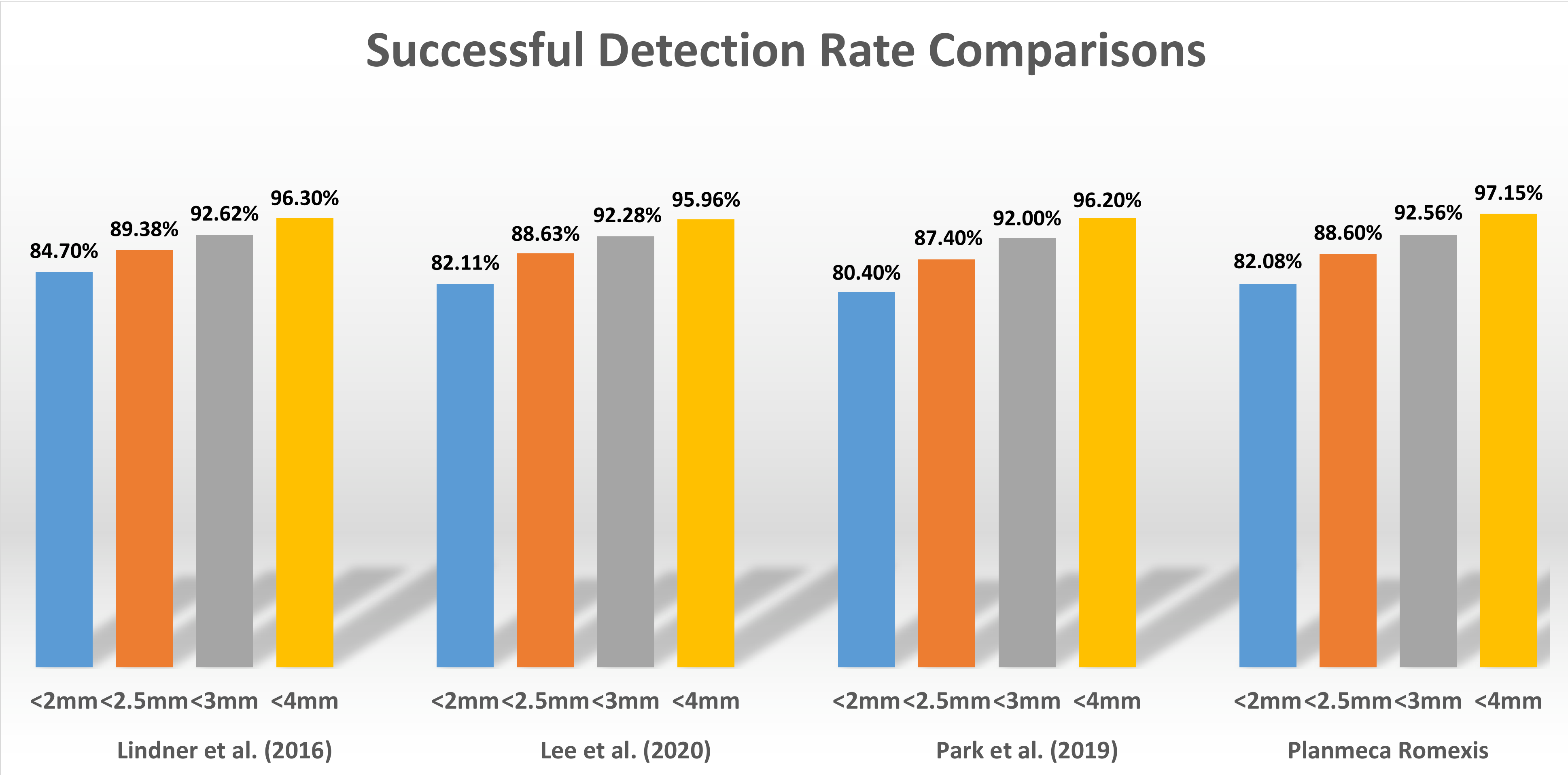
5.2 Comparison of hand tracing and computerised cephalometric analysis

5.2.1 Successful Detection Rate

Successful detection rates (SDR) for 2.0mm, 2.5mm, 3.0mm and 4.0mm precision ranges in this study are comparable with the three highest accuracy results for these SDRs reported in the literature to date by Lindner et al 2016,[83] Park et al. 2019[39] and Lee et al. 2020.[79] The comparison of these results is shown in Figure 5.1. The mean error for this study was 1.25 ± 0.91mm which is comparable with the reported mean error of expert landmarking of 1.26mm.[3, 18] The mean error reported by Lindner et al. 2016 was 1.2 ± 0.6mm[83] and by Lee et al. 2020 was 1.53 ± 1.74.[79] Park et al. 2019 did not report the mean error.[39]

The most significant and most widely reported result in the literature is the successful detection rate (SDR) for the 2.0mm precision range, as it is considered an acceptable level of error.[21] The SDR <2mm for this study was 82.08%. This is comparable with the highest results for this range reported in the literature, which are outlined in Table 1.5. It is worth noting that these studies were assessing accuracy of novel computer software programmes using a variety of AI-based methods for cephalometric landmark identification, which are not commercially available. Anuwongnukroh et al. (2017) carried out a study assessing the accuracy of automatic landmark identification with Carestream Dental, version 6.14, compared to manual tracing.[43] This software programme is commercially available. They found statistically significant differences in eight out of 13 landmarks and concluded that this automatic software programme was unable to compete with manual landmark identification.[43]

Figure 5.1 Comparison of the Successful Detection Rate (SDR) for 2.0mm, 2.5mm, 3.0mm and 4.0mm precision ranges of this study using Planmeca Romexis® (Version 5.1.1.R) dental imaging software with the three highest accuracy results reported for these precision ranges in the literature



5.2.2 Individual Landmark Accuracy

The most accurately located landmark in this study was the upper incisor tip, which had a successful detection rate (SDR) of 100% within 2mm. This is supported by the literature as the upper incisor tip is a landmark that is not generally associated with errors.[3]

Two out of the 16 landmarks identified in this study were significant outliers: A-point, with an SDR within 2mm of 57.35% and gonion with an SDR of 26.47% within 2mm. Both A-point and gonion are landmarks that are generally associated with subjectivity in identification but the recorded accuracy, particularly for gonion, in this study was lower than expected. As gonion is a constructed point, it is more frequently associated with errors than other landmarks.[3] Planmeca Romexis® (Version 5.1.1.R) dental imaging software struggled to identify gonion successfully and the mean error for this landmark was 3.24 ± 1.86mm. As demonstrated in the boxplots in Figures 4.8 and 4.9, gonion had increased variation in both the horizontal and vertical directions. This is consistent with the shape of the envelope of error for gonion reported in the literature.[3]

The literature has shown that the location of A-point can be subjective as it is located on a curve[3] and, as expected, it had greater variation in the vertical direction than in the horizontal direction as illustrated in the Bland-Altman plots in figures 4.3 and 4.4 and in the boxplots in Figures 4.8 and 4.9.

The successful detection rate (SDR) within 2mm for upper incisor apex and lower incisor apex was 77.94%. There was variation in both the horizontal and vertical directions which is consistent with the envelopes of error reported in the literature for these landmarks.[3]

5.2.3 Angular and Linear Measurements

Errors in landmark identification are the most common problem associated with lateral cephalometric analysis and, if the identification error is severe or if there are multiple errors, this can have a significant effect on the angular and linear measurements.[6] Given the low successful detection rate of gonion in this study, increased errors in MMPA and FMPA measurements might have been expected. The proportion of automated measurements within 2° of the hand traced measurements for FMPA was 82.35% and for MMPA was 88.76%, both of which indicate very good accuracy. The reason for the minimal effect of erroneous gonion positioning on these measurements relates to the distance between gonion and menton in the construction of the mandibular plane. If gonion and menton were closer together, the effect of the errors in landmark identification would be more significant.[6] While an improvement in the accuracy of the identification of gonion by the computer software programme would be ideal, the errors are not significantly effecting the accuracy of the angular measurements associated with gonion.

Overall, 80.98% of the automated angular and linear measurements were recorded within 2° or 2mm of the hand traced measurements. Four of the angles that recorded reduced levels of accuracy within 2° were associated with the position of the incisors: upper incisor to maxillary plane angle (54.41%), lower incisor to mandibular plane angle (55.88%), interincisal angle (57.35%) and lower incisor to occlusal plane angle (50%). This can be attributed to the errors associated with the identification of lower incisor apex and upper incisor apex, which would have a significant effect on the measurement of these angles. This is because upper incisor apex is relatively close to upper incisor tip and lower incisor apex is relatively close to lower incisor tip and therefore deviations in the positions of these landmarks will impact the long axis line of the affected incisor.[3] It would be important for the operator to check the position of upper incisor apex and lower incisor apex before performing an analysis. The accuracy of lower incisor to occlusal plane angle could also be affected by the angulation of the occlusal plane, which can be subjective and can vary between operators.

A study by Meriç and Naoumova (2020) compared the accuracy of fully automated, semi-automated, app-aided and manual cephalometric tracings.[102] The fully automated software programme used was CephX™, which is the same company that have partnered with Planmeca Romexis® to develop their automated cephalometric analysis software. Their study found that the fully automated analysis with CephX™ was the least accurate method and reported issues with accuracy of measurements involving gonion and the incisors, however they did not specifically measure the accuracy of landmark identification.[102]
In this study, while the accuracy for localisation of gonion, A-point, UIA and LIA is reduced, our findings suggest that the overall accuracy of cephalometric analysis with Planmeca Romexis® dental imaging software is similar to traditional hand tracing.

The null hypothesis for this study was rejected as the results show that there is a difference between automated cephalometric analysis with Planmeca Romexis® dental imaging software and traditional hand tracing. However, the differences measured between the two methods are clinically acceptable as they are within the acceptable measurement error range and the successful detection rate is within the inter-rater variability between two clinical experts.[83] Therefore, automated cephalometric analysis with Planmeca Romexis® dental imaging software is suitable for clinical use.

5.3 Implications for Practice

Traditional hand tracing of lateral cephalograms is time consuming and associated with errors, particularly in relation to landmark identification.[3, 7] Automated cephalometric analysis has many potential benefits if it is accurate and reliable. These include a reduction in the time taken to obtain an analysis, improved accuracy and reliability of the analysis and a reduction in operator errors due to clinicians' subjectivity.[8, 13]

The time taken to complete a cephalometric analysis by traditional hand tracing depends on several factors including operator experience, the type of analysis being used and how detailed the measurements are. Some clinicians do not perform measurements but simply visually assess the radiograph to examine the overall clinical picture.[52] Other clinicians perform a detailed cephalometric analysis, which is more accurate but time-consuming. There are issues with both of these approaches in terms of inefficiency, time commitment, consistency and likelihood of errors. The use of an accurate and reliable automated cephalometric analysis software programme has the potential to be of benefit for orthodontic practice due to its increased efficiency and therefore substantial time saving.

A recent study by Hwang et al. (2019) an AI-based software programme to human experts and found that both methods were equally accurate in terms of landmark identification.[96] They also found that the AI-based programme consistently identified the landmarks in identical positions when performing repeated measurements. Therefore, the AI program could be more reliable than human experts as there will always be some intra-examiner variability between humans.[96] This suggests that AI-based automated cephalometric analysis software programmes may have the potential to improve reliability and accuracy of cephalometric analysis.

5.4 Planmeca Romexis® (Version 5.1.1.R) Dental Imaging Software

Overall, the results of this study demonstrate reliability of Planmeca Romexis® (Version 5.1.1.R) dental imaging software for clinical use as the successful detection rate (SDR) <2mm of anatomical landmarks is 82.08% with a mean error of 1.25 ± 0.91mm, which is within the inter-rater reliability between two clinical experts. However, it is important that landmarks are checked by the operator before performing an analysis; particularly gonion, A-point, upper incisor apex and lower incisor apex.

The automatic angular and linear measurements were found to have an accuracy of 80.98% within 2° or 2mm respectively, which indicates a good level of agreement. However, the angles associated with the incisors had reduced accuracy and this necessitates the confirmation of upper incisor apex and lower incisor apex positions prior to performing an analysis.

5.5 Strengths and Limitations

5.5.1 Strengths

The strengths of this study were related to the study design and methodology. The study was sufficiently powered to meet the study aim and objectives. An inter-examiner reliability investigation was carried out to improve the accuracy and reliability of the results when comparing Planmeca Romexis® (Version 5.1.1.R) dental imaging software to the principal investigator. Intra-examiner reliability tracings were performed at an interval of four weeks to minimise recall bias. The study was thorough in that it measured the accuracy of both landmark identification and angular and linear measurements.

The findings of this study will help to inform clinicians who wish to use Planmeca Romexis® (Version 5.1.1.R) dental imaging software for automated cephalometric analysis about the strengths and weaknesses of the programme and the landmarks to check before performing an analysis.

5.5.2 Limitations

The limitations of this study relate to the gold standard chosen, measurement error and the types of analyses chosen. As the principal investigator is considered the gold standard, the best that the software program can achieve is to be equal to the gold standard i.e., comparable with a human expert. It is not possible to conclude that the software programme is superior to a human expert. If it was possible, a different gold standard could be used in future research to allow a more accurate assessment of the software programme's capability.

As the gold standard was hand tracing, the measurements in this study were carried out by hand using an orthodontic protractor/ruler and digital callipers. These methods are at risk of both measurement error and human error. It is also only possible to measure to 2 decimal places with the digital callipers and to measure to 0.5° or 0.5mm with the protractor/ruler. Digital measurements were not feasible in this study, but they would reduce the risk of error and give more precise measurements to more decimal places.

The x- and y-axes used for the landmark measurements were constructed parallel to the edges of the printed x-ray film through the centre of the ear rod. This method is easily reproducible for all the x-ray films, but it does not account for variation in head position and therefore the x-axis is not always perfectly parallel with the Frankfort plane and consequently the y-axis is not always perfectly perpendicular to the Frankfort plane. However, for radiographs where the head position was not perfect, the angle created between the constructed x-axis and the Frankfort plane is very small and therefore the effect on the x- and y-axis measurements is likely to be negligible.

No soft tissue landmarks or analysis was included in this study. It was not a deliberate decision to exclude soft tissue analysis, but the Eastman and the Downs analysis selected for the study only had hard tissue landmarks and measurements. The accuracy of soft tissue landmarking, and analysis should be examined in future research.

The hand tracings should ideally be made by a clinician who is independent of the study to minimise bias. However, this was not practical given the time commitment involved in completing the required tracings and subsequent measurements.

5.6 Future Research

The following could be examined in future research:

- The effect of digital image enhancement on accuracy of automated cephalometric analysis
- Comparison of automated cephalometric analysis with semi-automatic cephalometric analysis to facilitate computerised measurement of differences. This would reduce the risk of measurement error and minimise human error.
- Accuracy of automated cephalometric analysis could be measured with a custom analysis that also involves the most frequently used hard and soft tissue landmarks and angular and linear measurements
- The accuracy of other commercially available automated cephalometric analysis software programmes should be examined and reported in the scientific literature

Chapter 6 Conclusions

The findings of this study suggest that the accuracy of automated landmark identification with Planmeca Romexis® (Version 5.1.1.R) is comparable with the inter-rater reliability between two clinical experts with a successful detection rate (SDR) <2mm of 82.08% and a mean error of 1.25±0.91mm. Landmarks with lower accuracy levels were gonion, A-point, upper incisor apex and lower incisor apex and it is important for clinicians to be aware of potential inaccuracies arising from these in the analysis if using this software.

The findings of this study suggest that the accuracy of automated angular and linear measurements with Planmeca Romexis® (Version 5.1.1.R) is also comparable with the inter-rater reliability between two clinical experts with 80.98% of measurements recorded within 2° or 2mm of the hand traced measurements. The angular measurements involving the incisors were the least accurate, owing to the variation in accuracy of upper incisor apex and lower incisor apex.

Overall, the findings of this study are comparable with the highest performing artificial intelligence-based software programmes reported in the literature to date.

Automated cephalometric analysis with Planmeca Romexis® (Version 5.1.1.R) dental imaging software is suitable for clinical use, however, more accurate measurements may be possible if landmark locations are checked and amended if necessary prior to performing an analysis.

Chapter 7 Bibliography

1. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
2. Broadbent BH. A New X-Ray Technique and its Application to Orthodontia. The Angle Orthodontist. 1931;1(2):45-66.
3. Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. American journal of orthodontics. 1971;60(2):111-27.
4. Ricketts RM. Perspectives in the clinical application of cephalometrics. The first fifty years. Angle Orthod. 1981;51(2):115-50.
5. Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. Angle Orthod. 2008;78(1):145-51.
6. Baumrind S, Frantz RC. The reliability of head film measurements. 2. Conventional angular and linear measures. American journal of orthodontics. 1971;60(5):505-17.
7. Houston WJ, Maher RE, McElroy D, Sherriff M. Sources of error in measurements from cephalometric radiographs. Eur J Orthod. 1986;8(3):149-51.
8. Sommer T, Ciesielski R, Erbersdobler J, Orthuber W, Fischer-Brandies H. Precision of cephalometric analysis via fully and semiautomatic evaluation of digital lateral cephalographs. Dentomaxillofac Radiol. 2009;38(6):401-6.
9. Forsyth DB, Shaw WC, Richmond S. Digital imaging of cephalometric radiography, Part 1: Advantages and limitations of digital imaging. Angle Orthod. 1996;66(1):37-42.
10. Wenzel A, Gotfredsen E. Digital radiography for the orthodontist. Am J Orthod Dentofacial Orthop. 2002;121(2):231-5; quiz 192.
11. Visser H, Rodig T, Hermann KP. Dose reduction by direct-digital cephalometric radiography. Angle Orthod. 2001;71(3):159-63.
12. Forsyth DB, Davis DN. Assessment of an automated cephalometric analysis system. Eur J Orthod. 1996;18(5):471-8.
13. Shahidi S, Oshagh M, Gozin F, Salehi P, Danaei SM. Accuracy of computerized automatic identification of cephalometric landmarks by a designed software. Dentomaxillofac Radiol. 2013;42(1):20110187.
14. Broch J, Slagsvold O, Røsler M. Error in landmark identification in lateral radiographic headplates. Eur J Orthod. 1981;3(1):9-13.
15. Chen YJ, Chen SK, Yao JC, Chang HF. The effects of differences in landmark identification on the cephalometric measurements in traditional versus digitized cephalometry. Angle Orthod. 2004;74(2):155-61.
16. Houston WJ. The analysis of errors in orthodontic measurements. American journal of orthodontics. 1983;83(5):382-90.
17. Cohen AM, Linney AD. A low cost system for computer-based cephalometric analysis. Br J Orthod. 1986;13(2):105-8.
18. Parthasarathy S, Nugent ST, Gregson PG, Fay DF. Automatic landmarking of cephalograms. Comput Biomed Res. 1989;22(3):248-69.
19. Vucinić P, Trpovski Z, Sćepan I. Automatic landmarking of cephalograms using active appearance models. Eur J Orthod. 2010;32(3):233-41.
20. Gravely JF, Benzies PM. The clinical significance of tracing error in cephalometry. Br J Orthod. 1974;1(3):95-101.
21. Rakosi T. An Atlas and Manual of Cephalometric Radiography. London: Wolfe Medical Publications; 1982.

22. Liu JK, Chen YT, Cheng KS. Accuracy of computerized automatic identification of cephalometric landmarks. Am J Orthod Dentofacial Orthop. 2000;118(5):535-40.
23. Roden-Johnson D, English J, Gallerano R. Comparison of hand-traced and computerized cephalograms: landmark identification, measurement, and superimposition accuracy. Am J Orthod Dentofacial Orthop. 2008;133(4):556-64.
24. Ahlqvist J, Eliasson S, Welander U. The effect of projection errors on cephalometric length measurements. Eur J Orthod. 1986;8(3):141-8.
25. Ahlqvist J, Eliasson S, Welander U. The effect of projection errors on angular measurements in cephalometry. Eur J Orthod. 1988;10(4):353-61.
26. Richardson A. An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. American journal of orthodontics. 1966;52(9):637-51.
27. Midtgård J, Björk G, Linder-Aronson S. Reproducibility of cephalometric landmarks and errors of measurements of cephalometric cranial distances. Angle Orthod. 1974;44(1):56-61.
28. Bruks A, Enberg K, Nordqvist I, Hansson AS, Jansson L, Svenson B. Radiographic examinations as an aid to orthodontic diagnosis and treatment planning. Swed Dent J. 1999;23(2-3):77-85.
29. Nijkamp PG, Habets LLMH, Aartman IHA, Zentner A. The influence of cephalometrics on orthodontic treatment planning. The European Journal of Orthodontics. 2008;30(6):630-5.
30. Han UK, Vig KW, Weintraub JA, Vig PS, Kowalski CJ. Consistency of orthodontic treatment decisions relative to diagnostic records. Am J Orthod Dentofacial Orthop. 1991;100(3):212-9.
31. Devereux L, Moles D, Cunningham SJ, McKnight M. How important are lateral cephalometric radiographs in orthodontic treatment planning? Am J Orthod Dentofacial Orthop. 2011;139(2):e175-81.
32. Isaacson K, Thorn A, Atack N, Horner K, Whaites E. Guidelines for the use of radiographs in clinical orthodontics. 2015.
33. Kamoen A, Dermaut L, Verbeeck R. The clinical significance of error measurement in the interpretation of treatment results. Eur J Orthod. 2001;23(5):569-78.
34. Chen SK, Chen YJ, Yao CC, Chang HF. Enhanced speed and precision of measurement in a computer-assisted digital cephalometric analysis system. Angle Orthod. 2004;74(4):501-7.
35. Cohen AM, Ip HH, Linney AD. A preliminary study of computer recognition and identification of skeletal landmarks as a new method of cephalometric analysis. Br J Orthod. 1984;11(3):143-54.
36. Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. Eur J Orthod. 2000;22(5):499-508.
37. Polat-Ozsoy O, Gokcelik A, Toygar Memikoglu TU. Differences in cephalometric measurements: a comparison of digital versus hand-tracing methods. Eur J Orthod. 2009;31(3):254-9.
38. Goracci C, Ferrari M. Reproducibility of measurements in tablet-assisted, PC-aided, and manual cephalometric analysis. Angle Orthod. 2014;84(3):437-42.
39. Park JH, Hwang HW, Moon JH, Yu Y, Kim H, Her SB, et al. Automated identification of cephalometric landmarks: Part 1-Comparisons between the latest deep-learning methods YOLOV3 and SSD. Angle Orthod. 2019;89(6):903-9.

40. Wang CW, Huang CT, Hsieh MC, Li CH, Chang SW, Li WC, et al. Evaluation and Comparison of Anatomical Landmark Detection Methods for Cephalometric X-Ray Images: A Grand Challenge. IEEE Trans Med Imaging. 2015;34(9):1890-900.
41. Lindner C, Cootes TF, editors. Fully automatic cephalometric evaluation using random forest regression-voting. IEEE International Symposium on Biomedical Imaging; 2015: Citeseer.
42. Ibragimov B, Likar B, Pernus F, Vrtovec T, editors. Computerized cephalometry by game theory with shape-and appearance-based landmark refinement. Proceedings of International Symposium on Biomedical imaging (ISBI); 2015.
43. Anuwongnukroh N, Dechkunakorn S, Damrongsri S, Nilwarat C, Pudpong N, Radomsutthisarn W, et al., editors. Accuracy of Automatic Cephalometric Software on Landmark Identification2017: Institute of Physics Publishing.
44. Mohammad-Rahimi H, Nadimi M, Rohban MH, Shamsoddin E, Lee VY, Motamedian SR. Machine learning and orthodontics, current trends and the future opportunities: A scoping review. Am J Orthod Dentofacial Orthop. 2021.
45. Hung K, Montalvao C, Tanaka R, Kawai T, Bornstein MM. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. Dentomaxillofac Radiol. 2020;49(1):20190107.
46. Khanagar SB, Al-Ehaideb A, Vishwanathaiah S, Maganur PC, Patil S, Naik S, et al. Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making - A systematic review. J Dent Sci. 2021;16(1):482-92.
47. Lévy-Mandel AD, Venetsanopoulos AN, Tsotsos JK. Knowledge-based landmarking of cephalograms. Comput Biomed Res. 1986;19(3):282-309.
48. Tong W, Nugent S, Jensen G, Fay D, editors. An algorithm for locating landmarks on dental X-rays. Images of the Twenty-First Century Proceedings of the Annual International Engineering in Medicine and Biology Society; 1989: IEEE.
49. Douglas TS. Image processing for craniofacial landmark identification and measurement: a review of photogrammetry and cephalometry. Comput Med Imaging Graph. 2004;28(7):401-9.
50. Ren J, Liu D, Feng D, Shao J, Zhao R, Liao Y, et al., editors. A knowledge-based automatic cephalometric analysis method. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Vol 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat No 98CH36286); 1998: IEEE.
51. Cardillo J, Sid-Ahmed MA. An image processing system for locating craniofacial landmarks. IEEE Trans Med Imaging. 1994;13(2):275-89.
52. Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. Am J Orthod Dentofacial Orthop. 1998;113(2):173-9.
53. Saad A, El-Bialy A, Kandil A, Sayed A. Automatic cephalometric analysis using active appearance model and simulated annealing. ICGST Int J on Graphics, Vision and Image Processing, Special Issue on Image Retrieval and Representation. 2006;6:51-67.
54. Innes A, Ciesielski V, Mamutil J, John S, editors. Landmark detection for cephalometric radiology images using pulse coupled neural networks. Proc Int Conf on Artificial Intelligence; 2002: Citeseer.
55. Noble WS. What is a support vector machine? Nature Biotechnology. 2006;24(12):1565-7.

56. Chakrabartty S, Yagi M, Shibata T, Cauwenberghs G, editors. Robust cephalometric identification using support vector machines. Conference proceedings: ICME '03, International Conference on Multimedia and Expo; 2003.
57. Ciesielski V, Innes A, John S, Mamutil J. Genetic programming for landmark detection in cephalometric radiology images. International Journal of Knowledge based intelligent engineering Systems. 2003;7(3):164-71.
58. Zadeh LA. Information and control. Fuzzy sets. 1965;8(3):338-53.
59. Zadeh LA. Making computers think like people [fuzzy set theory]. IEEE spectrum. 1984;21(8):26-32.
60. El-Feghi I, Sid-Ahmed MA, Ahmadi M. Automatic localization of craniofacial landmarks for assisted cephalometry. Pattern Recognition. 2004;37(3):609-21.
61. Grau V, Alcañiz M, Juan MC, Monserrat C, Knoll C. Automatic localization of cephalometric Landmarks. J Biomed Inform. 2001;34(3):146-56.
62. Yue W, Yin D, Li C, Wang G, Xu T. Automated 2-D cephalometric analysis on X-ray images by a model-based approach. IEEE Trans Biomed Eng. 2006;53(8):1615-23.
63. Favaedi L, Petrou M. Cephalometric landmarks identification using probabilistic relaxation. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:4391-4.
64. Yang J, Ling X, Lu Y, Wei M, Ding G. Cephalometric image analysis and measurement for orthognathic surgery. Med Biol Eng Comput. 2001;39(3):279-84.
65. Mohseni H, Kasaei S, editors. Automatic localization of cephalometric landmarks2007.
66. Giordano D, Leonardi R, Maiorana F, Cristaldi G, Distefano ML, editors. Automatic landmarking of cephalograms by cellular neural networks. Conference on Artificial Intelligence in Medicine in Europe; 2005: Springer.
67. Kafieh R, Mehri A, Sadri S, editors. Automatic landmark detection in cephalometry using a modified active shape model with sub image matching. 2007 International Conference on Machine Vision; 2007: IEEE.
68. Rueda S, Alcañiz M. An approach for the automatic cephalometric landmark detection using mathematical morphology and active appearance models. Med Image Comput Comput Assist Interv. 2006;9(Pt 1):159-66.
69. Kaur A, Singh C. Automatic cephalometric landmark detection using Zernike moments and template matching. Signal, Image and Video Processing. 2015;9(1):117-32.
70. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017;69:S36-S40.
71. Ramesh A, Kambhampati C, Monson J, Drew P. Artificial intelligence in medicine. Annals of The Royal College of Surgeons of England. 2004;86(5):334-8.
72. Wong SH, Al-Hasani H, Alam Z, Alam A. Artificial intelligence in radiology: how will we be affected? European Radiology. 2019;29(1):141-3.
73. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019;380(14):1347-58.
74. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. Japanese Journal of Radiology. 2018;36(4):257-72.
75. Kim P. Convolutional neural network. MATLAB deep learning: Springer; 2017. p. 121-47.
76. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.
77. Kingsford C, Salzberg SL. What are decision trees? Nature Biotechnology. 2008;26(9):1011-3.

78. Wong T-T, Yeh P-Y. Reliable Accuracy Estimates from k-Fold Cross Validation. IEEE Transactions on Knowledge and Data Engineering. 2020;32(8):1586-94.
79. Lee JH, Yu HJ, Kim MJ, Kim JW, Choi J. Automated cephalometric landmark detection with confidence regions using Bayesian convolutional neural networks. BMC Oral Health. 2020;20(1):270.
80. Leonardi R, Giordano D, Maiorana F. An evaluation of cellular neural networks for the automatic identification of cephalometric landmarks on digital images. J Biomed Biotechnol. 2009;2009:717102.
81. Wang CW, Huang CT, Lee JH, Li CH, Chang SW, Siao MJ, et al. A benchmark for comparison of dental radiography analysis algorithms. Med Image Anal. 2016;31:63-76.
82. Ibragimov B, Likar B, Pernus F, Vrtovec T, editors. Automatic cephalometric X-ray landmark detection by applying game theory and random forests. Proc ISBI Int Symp on Biomedical Imaging; 2014.
83. Lindner C, Wang CW, Huang CT, Li CH, Chang SW, Cootes TF. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. Sci Rep. 2016;6:33581.
84. Chu C, Chen C, Nolte L, Zheng G. Fully automatic cephalometric x-ray landmark detection using random forest regression and sparse shape composition. submitted to Automatic Cephalometric X-ray Landmark Detection Challenge. 2014.
85. Arık S, Ibragimov B, Xing L. Fully automated quantitative cephalometry using convolutional neural networks. J Med Imaging (Bellingham). 2017;4(1):014501.
86. Wang S, Li H, Li J, Zhang Y, Zou B. Automatic Analysis of Lateral Cephalograms Based on Multiresolution Decision Tree Regression Voting. J Healthc Eng. 2018;2018:1797502.
87. Chen R, Ma Y, Chen N, Lee D, Wang W. Cephalometric Landmark Detection by Attentive Feature Pyramid Fusion and Regression-Voting. Springer International Publishing; 2019. p. 873-81.
88. Dai X, Zhao H, Liu T, Cao D, Xie L. Locating Anatomical Landmarks on 2D Lateral Cephalograms Through Adversarial Encoder-Decoder Networks. IEEE Access. 2019;7:132738-47.
89. Goutham END, Vasamsetti S, Kishore PVV, Sardana HK, editors. Automatic Localization of Landmarks in Cephalometric Images via Modified U-Net2019 2019: IEEE.
90. Qian J, Cheng M, Tao Y, Lin J, Lin H, editors. CephaNet: An Improved Faster R-CNN for Cephalometric Landmark Detection2019 2019: IEEE.
91. Song Y, Qiao X, Iwamoto Y, Chen Y-w. Automatic cephalometric landmark detection on X-ray images using a deep-learning method. Applied Sciences. 2020;10(7):2547.
92. Zhong Z, Li J, Zhang Z, Jiao Z, Gao X. An Attention-Guided Deep Regression Model for Landmark Detection in Cephalograms. Springer International Publishing; 2019. p. 540-8.
93. Gilmour L, Ray N, editors. Locating Cephalometric X-Ray Landmarks with Foveated Pyramid Attention. Medical Imaging with Deep Learning; 2020: PMLR.
94. Oh K, Oh IS, Le VNT, Lee DW. Deep Anatomical Context Feature Learning for Cephalometric Landmark Detection. IEEE J Biomed Health Inform. 2021;25(3):806-17.
95. Kim H, Shim E, Park J, Kim YJ, Lee U, Kim Y. Web-based fully automated cephalometric analysis by deep learning. Comput Methods Programs Biomed. 2020;194:105513.

96. Hwang HW, Park JH, Moon JH, Yu Y, Kim H, Her SB, et al. Automated identification of cephalometric landmarks: Part 2- Might it be better than human? Angle Orthod. 2020;90(1):69-76.
97. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. J Orofac Orthop. 2020;81(1):52-68.
98. Nishimoto S, Sotsuka Y, Kawai K, Ishise H, Kakibuchi M. Personal Computer-Based Cephalometric Landmark Detection With Deep Learning, Using Cephalograms on the Internet. J Craniofac Surg. 2019;30(1):91-5.
99. Anuwongnukroh N, Dechkunakorn S, Damrongsri S, Nilwarat C, Pudpong N, Radomsutthisarn W, et al. Assessment of the reliability of automatic cephalometric analysis software. Int J Mech Eng Robot Res. 2018;7(1):61-5.
100. Jeon S, Lee KC. Comparison of cephalometric measurements between conventional and automatic cephalometric analysis using convolutional neural network. Prog Orthod. 2021;22(1):14.
101. Planmeca and CephX Technologies LTD launch service for automatic cephalometric tracing and analysis [press release]. Finland: Planmeca Oy, March 23, 2017 2017.
102. Meriç P, Naoumova J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. Turk J Orthod. 2020;33(3):142-9.
103. Noothout JMH, De Vos BD, Wolterink JM, Postma EM, Smeets PAM, Takx RAP, et al. Deep Learning-Based Regression and Classification for Automatic Landmark Localization in Medical Images. IEEE Trans Med Imaging. 2020;39(12):4011-22.
104. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Archives of Orofacial Science. 2017;12(1).
105. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101-10.

Chapter 8 Appendices

8.1 Appendix 1: Literature Search Strategy

Search 1

(“Cephalometry/methods”[Mesh] OR “Cephalometry/standards”[Mesh] OR “Cephalometric analysis” OR Cephalometrics OR Cephalometry OR “Cephalometric measurement” OR “Lateral cephalogram”) AND (“Image Interpretation, Computer-Assisted”[Mesh] OR Computer-aided OR Computer-assisted OR Computerized OR Digital) AND (“Patterns Recognition, Automated”[Mesh] OR “Landmark identification” OR Landmarks) AND (Automatic OR “Fully automatic” OR Automated OR “Fully automated” OR “Manual tracing” OR “Hand tracing” OR Hand)

Search 2

(“Cephalometry/methods”[Mesh] OR “Cephalometry/standards”[Mesh] OR “Cephalometric analysis” OR Cephalometrics OR Cephalometry OR “Cephalometric measurement” OR “Lateral cephalogram”) AND (“Image Interpretation, Computer-Assisted”[Mesh] OR Computer-aided OR Computer-assisted OR Computerized OR Digital) AND (“Patterns Recognition, Automated”[Mesh] OR “Landmark identification” OR Landmarks) AND (Automatic OR “Fully automatic” OR Automated OR “Fully automated” OR “Automated Identification” OR “Manual tracing” OR “Hand tracing” OR Hand) AND (“Artificial Intelligence” OR “Deep Learning” OR “Machine learning”)

Search 3

(“Cephalometry/methods”[Mesh] OR “Cephalometry/standards”[Mesh] OR “Cephalometric analysis” OR Cephalometrics OR Cephalometry OR “Cephalometric measurement” OR “Lateral cephalogram”) AND (“Pattern Recognition, Automated”[Mesh] OR “Landmark identification” OR Landmarks) AND (Automatic OR “Fully automatic” OR Automated OR “Fully automated” OR “Automated Identification” OR “Manual tracing” OR “Hand tracing” OR Hand) AND (“Artificial Intelligence” OR “Intelligence, Artificial”[Mesh] OR “Deep Learning” OR “Machine learning”)

8.2 Appendix 2: Ethical Approval Letter

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Dr. Caroline Marron
Division of Public and Child Dental Health,
Dublin Dental University Hospital,
Lincoln Place,
Dublin 2,
D02 F859

4th February 2020

Ref: 191209

Title of Study: A comparison of hand tracing and computerised cephalometric analysis

Dear Dr. Marron,

Further to a meeting of the Faculty of Health Sciences Ethics Committee held in January. We are pleased to inform you that the above project has ethical approval to proceed.

This study has been ethically approved. We would advise you to seek review and comments on your DPIA from the DPO if required prior to study commencement'

As a researcher you must ensure that you comply with other relevant regulations, including DATA PROTECTION and HEALTH AND SAFETY.

Yours sincerely,

pp Fiona Brown

Prof. Jacintha O'Sullivan
Chairperson
Faculty Research Ethics Committee

Dámh na nEolaíochtaí Sláinte
Foirgneamh na Ceimice,
Coláiste na Tríonóide,
Ollscoil Átha Cliath,
Baile Átha Cliath 2, Éire.

Faculty of Health Sciences
Chemistry Building,
Trinity College Dublin,
The University of Dublin,
Dublin 2, Ireland.

www.healthsciences.tcd.ie

8.3 Appendix 3: Adult Participant Information Leaflet

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Adult Participant Information Leaflet

Name of Study: A comparison of hand tracing and computerised cephalometric analysis

Site	**HSE Orthodontic Department, SIMMS Building, Tallaght, Dublin 24** **Dublin Dental University Hospital, Lincoln Place, Dublin 2**
Principal Investigator(s) and Co-Investigator(s) (insert names, titles and contact details)	**Principal Investigator: Dr Caroline Marron** **Email: caroline.marron@dental.tcd.ie** **Contact number: 01 6127303**
Data Controllers	**Trinity College Dublin (for research data)** **HSE Orthodontic Department, SIMMS Building, Tallaght, Dublin 24 (for hospital medical records)**
Data Protection Officer	**Data Protection Officer** **Secretary's Office** **Trinity College Dublin** **Dublin 2** **Deputy Data Protection Officer Dublin Mid-Leinster: Debbie Keyes.**

Dear Parent/Guardian

My name is Dr Caroline Marron and I am a Clinical Doctoral Candidate in Orthodontics at the School of Dental Science, Trinity College Dublin. I would like to invite you to take part in a research study. Before you decide, you need to understand why the research is being done and what it would involve for you. Please take the time to read the following information carefully. Ask questions if anything you read is not clear or if you would like more information.

This leaflet has five main parts:

Part 1 – The Study
Part 2 – Data Protection
Part 3 – Costs, Funding and Approval
Part 4 – Future Research
Part 5 – Further Information

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 1 – The Study

Why is this study being done?

Lateral cephalograms are dental x-rays taken prior to having orthodontic treatment. These x-rays are measured after they have been taken and the measurements help to make the decision as to what type of orthodontic treatment patients require. Traditionally these measurements were done by hand but more recently, computer programs have been introduced that automatically measure the x-ray, which saves a lot of time. The accuracy of these computer programs has not been verified and this study is being done to see if the computer measurements are as accurate as the hand measurements.

Why have I been invited to take part?

You have been invited to take part because your orthodontist has prescribed one of these x-rays. We are recruiting 80 participants attending the HSE Orthodontic Department, SIMMS Building, Tallaght who, like you, have been prescribed an x-ray by their orthodontist.

Do I have to take part? Can I withdraw?

Participation in this study is voluntary. There are no advantages, disadvantages or consequences to your orthodontic treatment should you choose to participate or not. Consent can be withdrawn up to the point of anonymisation of the x-rays. After they have been anonymised they will no longer be identifiable and therefore cannot be withdrawn. If you wish to withdraw from the study please contact the principal researcher Dr Caroline Marron:

- Email: caroline.marron@dental.tcd.ie
- Contact number: 01 6127303

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

What happens if I change my mind?

You can change your mind at any time up to the point of anonymisation of the x-rays. If you choose not to continue to take part, this will not affect your orthodontic treatment in any way. If you choose not to take part any more, you will be asked to fill in a withdrawal form. If you withdraw your x-ray will not be used in the study however it will still be stored as part of your medical record in the HSE Orthodontic Department, SIMMS Building, Tallaght.

How will the study be carried out?

The study is taking place this year in the HSE Orthodontic Department, SIMMS Building, Tallaght. The radiographer will irreversibly anonymise a copy of your x-ray after it has been taken by removing all personal details from it. The original x-ray will be stored on your medical record and be used for the purposes for which it was prescribed. The anonymised copy of your x-ray will be given to the principal researcher, Dr Caroline Marron for analysis. Each x-ray will be measured first by hand and then using the computer program and a comparison will be made to see if the computer program measurements are as accurate as the hand measurements. There will be 80 patients participating in the study.

What will happen to me if I decide to take part?

If you decide to take part, a copy of your x-ray will be anonymised and given to the principal researcher for analysis. Your original x-ray will be stored, as usual, as part of your medical record in the HSE Orthodontic Department, SIMMS Building, Tallaght. The anonymised x-ray will be measured first by hand and then using the computer program. A comparison will be made to see if the computer program measurements are as accurate as the hand measurements. Your orthodontic care will proceed as normal. Once your prescribed x-ray has been taken there are no further obligations as part of the study and we will not contact you again.

What will happen to my Samples and Data?

All hard copy records (consent forms, anonymised x-rays) will be stored securely. Consent forms will be stored in a locked cabinet in a swipe access controlled room in the HSE Orthodontic Department, SIMMS Building, Tallaght, with access solely by the radiographer. The anonymised x-rays will be stored in a locked cabinet with access solely by the lead investigator. The cabinet is located in a locked room in Dublin Dental University Hospital with swipe card access only. These records will be kept for a period of 7 years and will then be destroyed.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Are there any benefits to taking part in this research?

Taking part in this study will not directly benefit you. However, research performed with your anonymised x-ray may help us to better understand if the x-ray measurement computer programs are accurate and reliable. By participating, you are helping to advance orthodontic care and technology for future generations.

Are there any risks to me or others if I take part?

There are no anticipated risks involved in partaking in this study. Your x-ray is being prescribed by your orthodontist independent of this study. Your x-ray will be irreversibly anonymised prior to being measured and therefore will not be identifiable.

Will I be told the outcome of the study? Will I be told the results of any tests or investigations performed as part of this study that relate to me?

The results of the study will be analysed and published in a Doctoral Thesis, which will be submitted to Dublin Dental University Hospital and Trinity College Dublin. The results of the study will be reported in orthodontic/scientific journals and disclosed at orthodontic/scientific conferences. These results will be used for educational and teaching purposes. No information which reveals your identity will be disclosed. Your x-ray will be interpreted independently by the prescribing orthodontist as part of your orthodontic diagnosis and treatment plan.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 2 – Data Protection

What information about me (personal data) will be used as part of this study? Will my medical records be accessed?

Your written consent form and your x-ray will used as part of this study. Your x-ray will be anonymised prior to being given to the principal investigator for analysis and therefore will not be identifiable.

What will happen to my personal data?

Your written consent form will be stored in a locked cabinet in a swipe access controlled room in the HSE Orthodontic Department, SIMMS Building, Tallaght, with access solely by the radiographer. Your anonymised x-ray will be stored in a locked cabinet with access solely by the lead investigator. The cabinet is located in a locked room in Dublin Dental University Hospital with swipe card access only. These records will be kept for a period of 7 years and will then be destroyed.

Who will access and use my personal data as part of this study?

The radiographer in the HSE Orthodontic Department, SIMMS Building, Tallaght will have access to your consent form. The principal investigator, Dr Caroline Marron will have access to your anonymised x-ray.

Will my personal data be kept confidential? How will my data be kept safe?

Your identity will remain confidential in this study. Your consent form will be stored securely as described above. Your x-ray will be irreversibly anonymised and therefore will not be identifiable. The researchers are also bound by a professional code of secrecy which would mean disciplinary action if confidentiality was breached. Training in data protection law has been carried out by the researchers.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

What is the lawful basis to use my personal data?

By law,[1] we can use your personal information for scientific research[2] (in the public interest[3]). We will also ask for your explicit consent to use your data as a requirement of the Irish Health Research Regulations.

What are my rights?

You are entitled to:

- The right to access to your data and receive a copy of it
- The right to restrict or object to processing of your data
- The right to object to any further processing of the information we hold about you (except where it is de-identified)
- The right to have inaccurate information about you corrected or deleted
- The right to receive your data in a portable format and to have it transferred to another data controller
- The right to request deletion of your data

By law you can exercise the following rights in relation to your personal data, unless the request would make it impossible or very difficult to conduct the research. You can exercise these rights by contacting your study Dr Caroline Marron, email: caroline.marron@dental.tcd.ie, contact number: 01 6127303 or the Trinity College Data Protection Officer, Secretary's Office, Trinity College Dublin, Dublin 2, Ireland. Email: dataprotection@tcd.ie. Website: www.tcd.ie/privacy.

[1] The European General Data Protection Regulation (GDPR)
[2] *Article 9(2) (j))*
[3] *(Article 6(1)(e)*

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 3 – Costs, Funding and Approval

Has this study been approved by a research ethics committee?

Yes, this study has been approved by the Trinity College Faculty of Health Sciences Research Ethics Committee. Approval was granted on 4th February 2020.

Who is organising and funding this study? Will the results be used for commercial purposes?

The study is being organised and funded by Dublin Dental University Hospital and Trinity College Dublin. The results will not be used for commercial purposes.

Is there any payment for taking part? Will it cost me anything if I agree to take part?

No, we are not paying patients to take part in the study.

Part 4 – Future Research

Will my personal data and/or biological material be used in future studies? (May not apply)

N/A

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 5 – Further Information

Who should I contact for information or complaints?

If you have any concerns or questions, you can contact:

- Principal Investigator: Dr Caroline Marron
    - Email: caroline.marron@dental.tcd.ie
    - Contact number: 01 6127303
- Deputy Data Protection Officer Dublin Mid-Leinster: Debbie Keyes. Email: ddpo.dml@hse.ie Contact number: 057-9357876/045-880496
- Data Protection Officer, Trinity College Dublin: Data Protection Officer, Secretary's Office, Trinity College Dublin, Dublin 2, Ireland. Email: dataprotection@tcd.ie. Website: www.tcd.ie/privacy.

Under GDPR, if you are not satisfied with how your data is being processed, you have the right to lodge a complaint with the Office of the Data Protection Commission, 21 Fitzwilliam Square South, Dublin 2, Ireland. Website: www.dataprotection.ie.

Will I be contacted again?

No, you will not be contacted again. If you would like to take part in this study, you will be asked to sign the Consent Form on the next page. You will be given a copy of this information leaflet to keep. Thank you for taking the time to consider participating in this study.

8.4 Appendix 4: Parent/Guardian Participant Information Leaflet

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Information Leaflet (Parent/Guardian)

Name of Study: A comparison of hand tracing and computerised cephalometric analysis

Site	**HSE Orthodontic Department, SIMMS Building, Tallaght, Dublin 24** **Dublin Dental University Hospital, Lincoln Place, Dublin 2**
Principal Investigator(s) and Co-Investigator(s) (insert names, titles and contact details)	**Principal Investigator: Dr Caroline Marron** **Email: caroline.marron@dental.tcd.ie** **Contact number: 01 6127303**
Data Controllers	**Trinity College Dublin (for research data)** **HSE Orthodontic Department, SIMMS Building, Tallaght, Dublin 24 (for hospital medical records)**
Data Protection Officer	**Data Protection Officer** **Secretary's Office** **Trinity College Dublin** **Dublin 2** **Deputy Data Protection Officer Dublin Mid-Leinster: Debbie Keyes.**

Dear Parent/Guardian

My name is Dr Caroline Marron and I am a Clinical Doctoral Candidate in Orthodontics at the School of Dental Science, Trinity College Dublin. I would like to invite your child to take part in a research study. Before you decide, you need to understand why the research is being done and what it would involve for you and your child. Please take the time to read the following information carefully. Ask questions if anything you read is not clear or if you would like more information.

This leaflet has five main parts:

Part 1 – The Study
Part 2 – Data Protection
Part 3 – Costs, Funding and Approval
Part 4 – Future Research
Part 5 – Further Information

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 1 – The Study

Why is this study being done?

Lateral cephalograms are dental x-rays taken prior to having orthodontic treatment. These x-rays are measured after they have been taken and the measurements help to make the decision as to what type of orthodontic treatment patients require. Traditionally these measurements were done by hand but more recently, computer programs have been introduced that automatically measure the x-ray, which saves a lot of time. The accuracy of these computer programs has not been verified and this study is being done to see if the computer measurements are as accurate as the hand measurements.

Why have I been invited to take part?

You have been invited to take part because your child's orthodontist has prescribed one of these x-rays. We are recruiting 80 participants attending the HSE Orthodontic Department, SIMMS Building, Tallaght who, like your child, have been prescribed an x-ray by their orthodontist.

Do I have to take part? Can I withdraw?

Participation in this study is voluntary. There are no advantages, disadvantages or consequences to your child's orthodontic treatment should they choose to participate or not. Consent can be withdrawn up to the point of anonymisation of the x-rays. After they have been anonymised they will no longer be identifiable and therefore cannot be withdrawn. If you wish to withdraw from the study please contact the principal researcher Dr Caroline Marron:

- Email: caroline.marron@dental.tcd.ie
- Contact number: 01 6127303

What happens if I change my mind?

You can change your mind at any time up to the point of anonymisation of the x-rays. If you choose not to continue to take part, this will not affect your child's orthodontic treatment in any way. If you choose not to take part any more, you will be asked to fill in a withdrawal form. If you withdraw your child's radiograph will not be used in the study however it will still be stored as part of their medical record in the HSE Orthodontic Department, SIMMS Building, Tallaght.

How will the study be carried out?

The study is taking place this year in the HSE Orthodontic Department, SIMMS Building, Tallaght. The radiographer will irreversibly anonymise a copy of your child's x-ray after it has been taken by removing all personal details from them. The original x-ray will be stored on your child's medical record and be used for the purposes for which it was prescribed. The anonymised copy of your child's x-ray will be given to the principal researcher, Dr Caroline Marron for analysis. Each x-ray will be measured first by hand and then using the computer program and a comparison will be made to see if the computer program measurements are as accurate as the hand measurements. There will be 80 patients participating in the study.

What will happen to me if I decide to take part?

If you decide to take part, a copy of your child's x-ray will be anonymised and given to the principal researcher for analysis. Their original x-ray will be stored, as usual, as part of their medical record in the HSE Orthodontic Department, SIMMS Building, Tallaght. The anonymised x-ray will be measured first by hand and then using the computer program. A comparison will be made to see if the computer program measurements are as accurate as the hand measurements. Your child's orthodontic care will proceed as normal. Once your child's prescribed x-ray has been taken there are no further obligations as part of the study and we will not contact you again.

What will happen to my Samples and Data?

All hard copy records (consent forms, anonymised x-rays) will be stored securely. Consent forms will be stored in a locked cabinet in a swipe access controlled room in the HSE Orthodontic Department, SIMMS Building, Tallaght, with access solely by the radiographer. The anonymised x-rays will be stored in a locked cabinet with access solely by the lead investigator. The cabinet is located in a locked room in Dublin Dental University Hospital with swipe card access only. These records will be kept for a period of 7 years and will then be destroyed.

Are there any benefits to taking part in this research?

Taking part in this study will not directly benefit you or your child. However, research performed with your child's anonymised x-ray may help us to better understand if the x-ray measurement computer programs are accurate and reliable. By participating, you are helping to advance orthodontic care and technology for future generations.

Are there any risks to me or others if I take part?

There are no anticipated risks involved in partaking in this study. Your child's radiograph is being prescribed by their orthodontist independent of this study. Your child's x-ray will be irreversibly anonymised prior to being measured and therefore will not be identifiable.

Will I be told the outcome of the study? Will I be told the results of any tests or investigations performed as part of this study that relate to me?

The results of the study will be analysed and published in a Doctoral Thesis, which will be submitted to Dublin Dental University Hospital and Trinity College Dublin. The results of the study will be reported in orthodontic/scientific journals and disclosed at orthodontic/scientific conferences. These results will be used for educational and teaching purposes. No information which reveals your child's identity will be disclosed. Your child's radiograph will be interpreted independently by the prescribing orthodontist as part of your orthodontic diagnosis and treatment plan.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 2 – Data Protection

What information about me (personal data) will be used as part of this study? Will my medical records be accessed?

Your written consent form and your child's x-ray will used as part of this study. Your child's x-ray will be anonymised prior to being given to the principal investigator for analysis and therefore will not be identifiable.

What will happen to my personal data?

Your written consent form will be stored in a locked cabinet in a swipe access controlled room in the HSE Orthodontic Department, SIMMS Building, Tallaght, with access solely by the radiographer. Your child's anonymised x-ray will be stored in a locked cabinet with access solely by the lead investigator. The cabinet is located in a locked room in Dublin Dental University Hospital with swipe card access only. These records will be kept for a period of 7 years and will then be destroyed.

Who will access and use my personal data as part of this study?

The radiographer in the HSE Orthodontic Department, SIMMS Building, Tallaght will have access to your consent form. The principal investigator, Dr Caroline Marron will have access to your child's anonymised x-ray.

Will my personal data be kept confidential? How will my data be kept safe?

Your identity and your child's identity will remain confidential in this study. Your consent form will be stored securely as described above. Your child's x-ray will be irreversibly anonymised and therefore will not be identifiable. The researchers are also bound by a professional code of secrecy which would mean disciplinary action if confidentiality was breached. Training in data protection law has been carried out by the researchers.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

What is the lawful basis to use my personal data?

By law,[1] we can use your personal information for scientific research[2] (in the public interest[3]). We will also ask for your explicit consent to use your data as a requirement of the Irish Health Research Regulations.

What are my rights?

You are entitled to:

- The right to access to your data and receive a copy of it
- The right to restrict or object to processing of your data
- The right to object to any further processing of the information we hold about you (except where it is de-identified)
- The right to have inaccurate information about you corrected or deleted
- The right to receive your data in a portable format and to have it transferred to another data controller
- The right to request deletion of your data

By law you can exercise the following rights in relation to your personal data, unless the request would make it impossible or very difficult to conduct the research. You can exercise these rights by contacting your study Dr Caroline Marron, email: caroline.marron@dental.tcd.ie, contact number: 01 6127303 or the Trinity College Data Protection Officer, Secretary's Office, Trinity College Dublin, Dublin 2, Ireland. Email: dataprotection@tcd.ie. Website: www.tcd.ie/privacy.

[1] The European General Data Protection Regulation (GDPR)
[2] *Article 9(2) (j))*
[3] *(Article 6(1)(e)*

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 3 – Costs, Funding and Approval

Has this study been approved by a research ethics committee?

Yes, this study has been approved by the Trinity College Faculty of Health Sciences Research Ethics Committee. Approval was granted on 4th February 2020.

Who is organising and funding this study? Will the results be used for commercial purposes?

The study is being organised and funded by Dublin Dental University Hospital and Trinity College Dublin. The results will not be used for commercial purposes.

Is there any payment for taking part? Will it cost me anything if I agree to take part?

No, we are not paying patients to take part in the study.

Part 4 – Future Research

Will my personal data and/or biological material be used in future studies? (May not apply)

N/A

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Part 5 – Further Information

Who should I contact for information or complaints?

If you have any concerns or questions, you can contact:

- Principal Investigator: Dr Caroline Marron
  - Email: caroline.marron@dental.tcd.ie
  - Contact number: 01 6127303
- Deputy Data Protection Officer Dublin Mid-Leinster: Debbie Keyes. Email: ddpo.dml@hse.ie Contact number: 057-9357876/045-880496
- Data Protection Officer, Trinity College Dublin: Data Protection Officer, Secretary's Office, Trinity College Dublin, Dublin 2, Ireland. Email: dataprotection@tcd.ie. Website: www.tcd.ie/privacy.

Under GDPR, if you are not satisfied with how your data is being processed, you have the right to lodge a complaint with the Office of the Data Protection Commission, 21 Fitzwilliam Square South, Dublin 2, Ireland. Website: www.dataprotection.ie.

Will I be contacted again?

No, you will not be contacted again. If you would like to take part in this study, you will be asked to sign the Consent Form on the next page. You will be given a copy of this information leaflet to keep. Thank you for taking the time to consider participating in this study.

8.5 Appendix 5: Child Participant Information Leaflet

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

OSPIDÉAL DÉADACH OLLSCOILE ÁTHA CLIATH
DUBLIN DENTAL UNIVERSITY HOSPITAL

CHILD PARTICIPANT INFORMATION LEAFLET

Title of study: A comparison of hand tracing and computerised cephalometric analysis.

Dear Orthodontic Patient,

I would like to invite you to take part in a research project. Please take the time to read the following information carefully and, if you have any question please ask!

WHO I AM AND WHAT THIS STUDY IS ABOUT
I am an orthodontic postgraduate here in the orthodontic department in Tallaght. This means that I am a dentist who is training to be an orthodontist! Your orthodontist has prescribed an x-ray for you today. Your x-ray will be measured by the orthodontist and used to help plan your orthodontic treatment. These x-rays are measured by hand using a ruler and a protractor but we now have this new computer program that does the measurements for us, which is great! But we have to check if the computer program is as accurate as the hand measurements before it's used for everyone. So I am hoping to find that out!

WHAT WILL TAKING PART INVOLVE?
If you agree to take part, after you have had your x-ray taken by the radiographer today I will measure it by hand and then measure it using the computer program and see if the measurements are the same. So I won't be doing anything directly to you, I will just be looking at your x-ray. There will be no changes to your orthodontic treatment as a result of taking part or not taking part in the study. The x-ray won't be identifiable as you when I am taking measurements from it.

WHY HAVE YOU BEEN INVITED TO TAKE PART?
You have been asked to take part because your orthodontist has prescribed an x-ray for you which would be useful for us to measure as part of this study. We will be asking 80 patients to take part in total.

DO YOU HAVE TO TAKE PART?
You do not have to take part in this study if you do not wish to and there will be no changes to your orthodontic treatment whether you take part or not.

WHAT DO YOU NEED TO DO TO TAKE PART IN THIS STUDY?
Your mum/dad will let us know if you're happy to take part.

Thank you for taking the time to read this information and don't forget to ask questions if you have any!

Yours faithfully,

Dr. Caroline Marron
B.A., B. Dent. Sc. (Hons), MFD RCSI., MDPH., PGDipHCM.
D. Ch. Dent (Orthodontics) Candidate

EHS 2019

8.6 Appendix 6: Adult Consent Form

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

OSPIDÉAL DÉADACH OLLSCOILE ÁTHA CLIATH
DUBLIN DENTAL UNIVERSITY HOSPITAL

PATIENT CONSENT FORM

STUDY NAME: A comparison of hand tracing and computerised cephalometric analysis

Please ask any questions you may have when reading each of the statements.
Thank you for participating.

Please tick the box if you agree with the statement. Please feel free to ask questions if there is something you do not understand.

General	**Tick box**
I confirm I have read and understood the **Information Leaflet** for the above study. The information has been fully explained to me and I have been able to ask questions, all of which have been answered to my satisfaction.	
I understand that this study **is entirely voluntary, and if I decide that I do not want to take part, I can stop taking part in this study up to the point of anonymization of the x-rays without giving a reason**. I understand that deciding not to take part will not affect my future orthodontic care.	
I understand that I **will not be paid for taking part in this study**.	
I know how to contact the research team if I need to.	
I agree to take part in this research study having been fully informed of the **risks and benefits** which are set out in full in the information leaflet which I have been provided with.	

Data processing	Tick box
I understand that the processing of personal data about me will be protected in accordance with the General Data Protection Regulation.	
I understand that there are **no direct benefits to me** from participating in this study. I understand that **results from analysis of my personal information as part of this study will not be given to me.**	

Patient Name (Block Capitals) Patient Signature Date

--- --

Witness Name (Block Capitals) Witness Signature Date

To be completed by the Principal Investigator or nominee.

I, the undersigned, have taken the time to fully explain to the above patient the nature and purpose of this study in a way that they could understand. I have explained the risks and possible benefits involved. I have invited them to ask questions on any aspect of the study that concerned them.
I have given a copy of the information leaflet and consent form to the participant with contacts of the study team

Researcher name: Dr. Caroline Marron

Title and Qualifications: Postgraduate in Orthodontics, B.A., B. Dent. Sc. (Hons)., MFD RCSI., MDPH., PGDipHCM.

Signature:

Date:

8.7 Appendix 7: Parent/Guardian Consent Form

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

PATIENT CONSENT FORM (PARENT/GUARDIAN)

STUDY NAME: A comparison of hand tracing and computerised cephalometric analysis

Please ask any questions you may have when reading each of the statements.

Thank you for participating.

Please tick the box if you agree with the statement. Please feel free to ask questions if there is something you do not understand.

General	**Tick box**
I confirm I have read and understood the **Information Leaflet** for the above study. The information has been fully explained to me and I have been able to ask questions, all of which have been answered to my satisfaction.	
I understand that this study **is entirely voluntary, and if I decide that I do not want my child to take part, I can stop them taking part in this study up to the point of anonymization of the x-rays without giving a reason**. I understand that deciding not to take part will not affect my child's future orthodontic care.	
I understand that I **will not be paid for taking part in this study**.	
I know how to contact the research team if I need to.	
I agree to take part in this research study having been fully informed of the **risks and benefits** which are set out in full in the information leaflet which I have been provided with.	

FHS 201

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Data processing	**Tick box**
I understand that the processing of personal data about me and my child will be protected in accordance with the General Data Protection Regulation.	
I understand that there are **no direct benefits to me or my child** from participating in this study. I understand that **results from analysis of my child's personal information as part of this study will not be given to me.**	

Parent/Guardian Name (Block Capitals)   Parent/Guardian Signature   Date

--   --

Witness Name (Block Capitals)   Witness Signature   Date

To be completed by the Principal Investigator or nominee.

I, the undersigned, have taken the time to fully explain to the above patient the nature and purpose of this study in a way that they could understand. I have explained the risks and possible benefits involved. I have invited them to ask questions on any aspect of the study that concerned them.

I have given a copy of the information leaflet and consent form to the participant with contacts of the study team

Researcher name: Dr. Caroline Marron
Title and Qualifications: Postgraduate in Orthodontics, B.A., B. Dent. Sc. (Hons)., MFD RCSI., MDPH., PGDipHCM.
Signature:
Date:

EHS 2019

8.8 Appendix 8: Bland-Altman Plots

Figure 8.1: Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for anterior nasal spine (ANS)

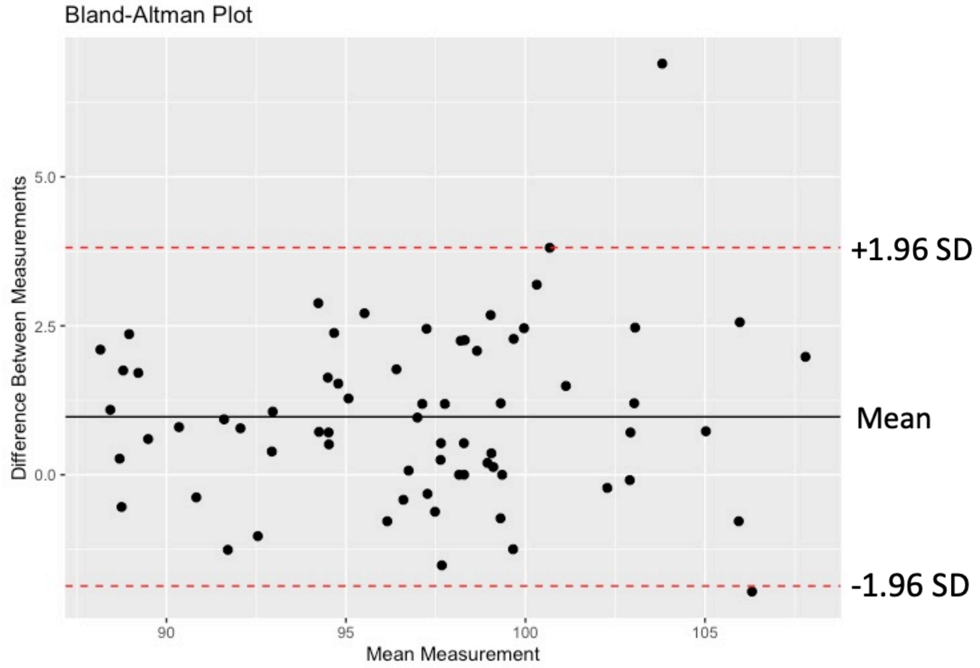
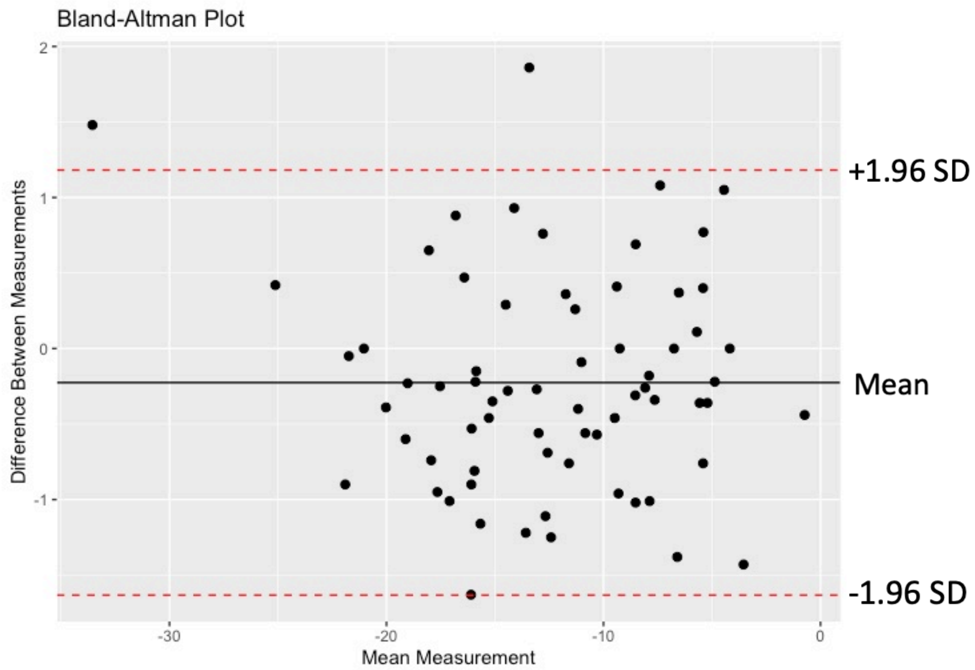


Figure 8.2: Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for anterior nasal spine (ANS)



Figures 8.1 and 8.2 show that the horizontal discrepancy was greater than the vertical discrepancy for the identification of ANS by hand tracing and automated tracing.

Figure 8.3 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for posterior nasal spine (PNS)

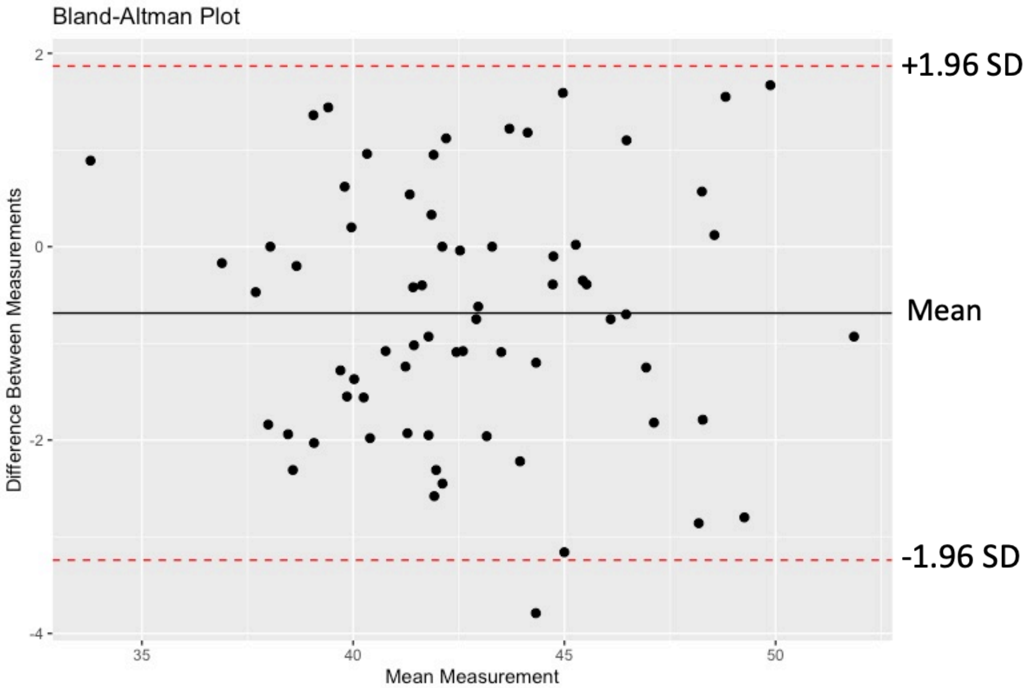


Figure 8.4 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for posterior nasal spine (PNS)

Figures 8.3 and 8.4 show that the horizontal discrepancy was greater than the vertical discrepancy for the identification of PNS by hand tracing and automated tracing.

Figure 8.5 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for B-Point

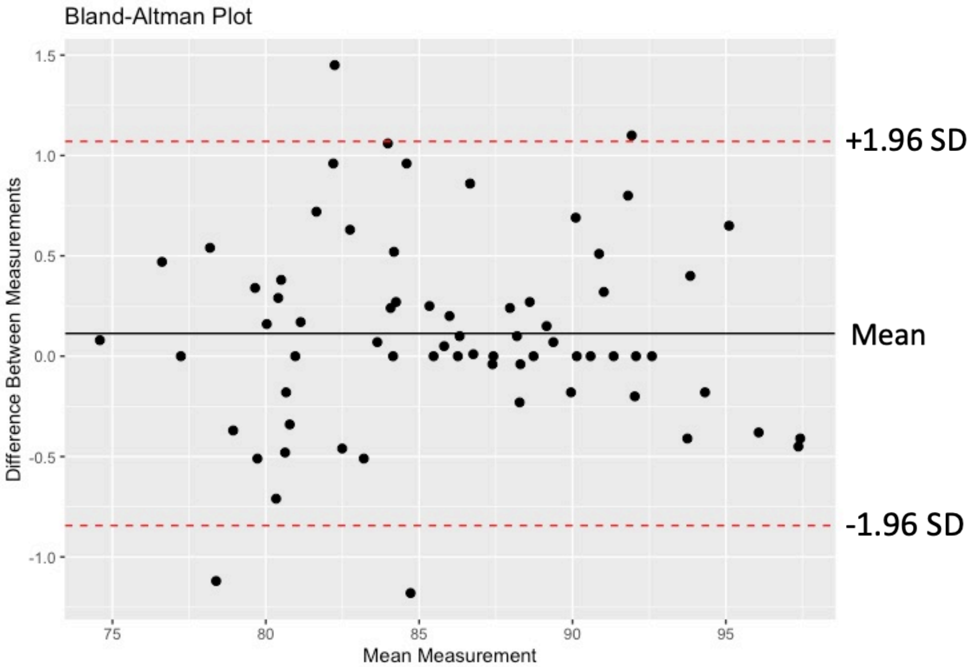
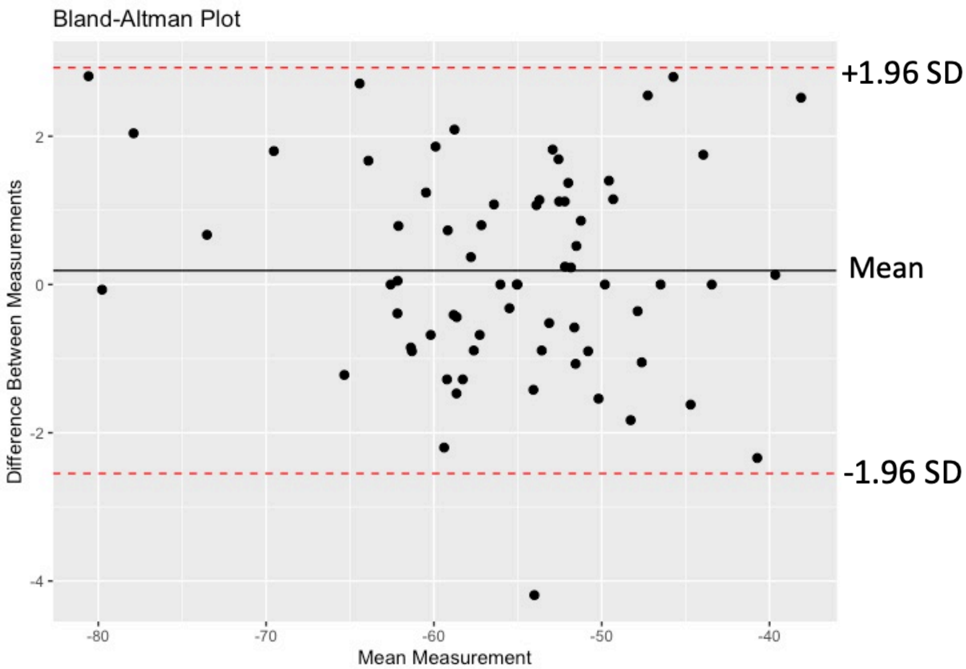


Figure 8.6 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for B-Point



Figures 8.5 and 8.6 show that the vertical discrepancy was greater than the horizontal discrepancy for the identification of B-Point by hand tracing and automated tracing.

Figure 8.7 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for upper incisor apex (UIA)

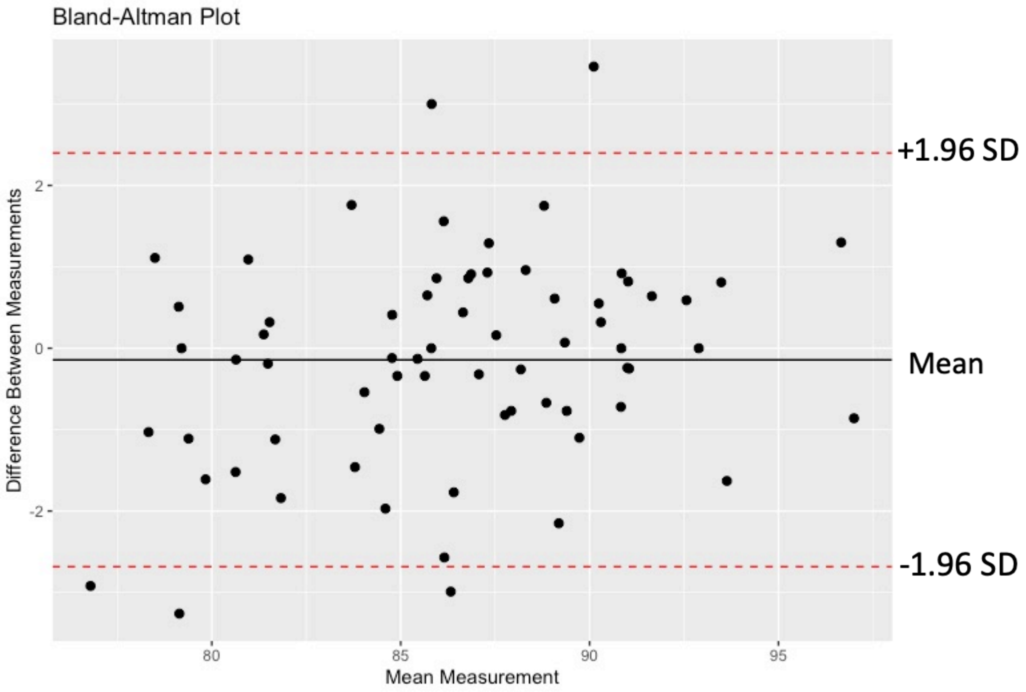
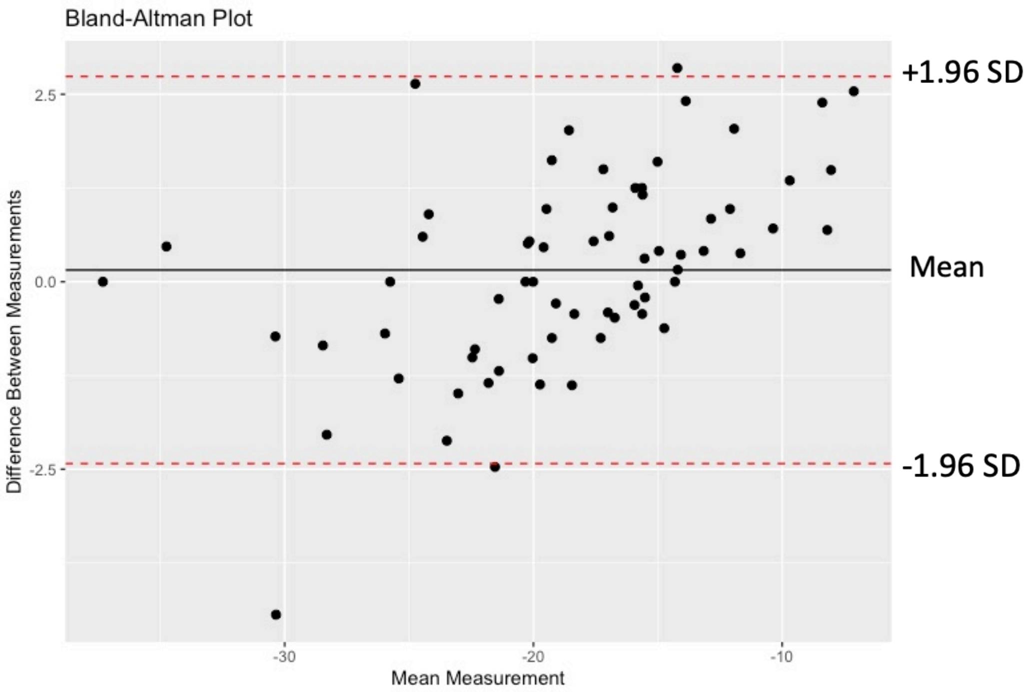


Figure 8.8 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for upper incisor apex (UIA)



Figures 8.7 and 8.8 show that the horizontal and vertical discrepancies were similar for the identification of upper incisor apex by hand tracing and automated tracing.

Figure 8.9 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for lower incisor apex (LIA)

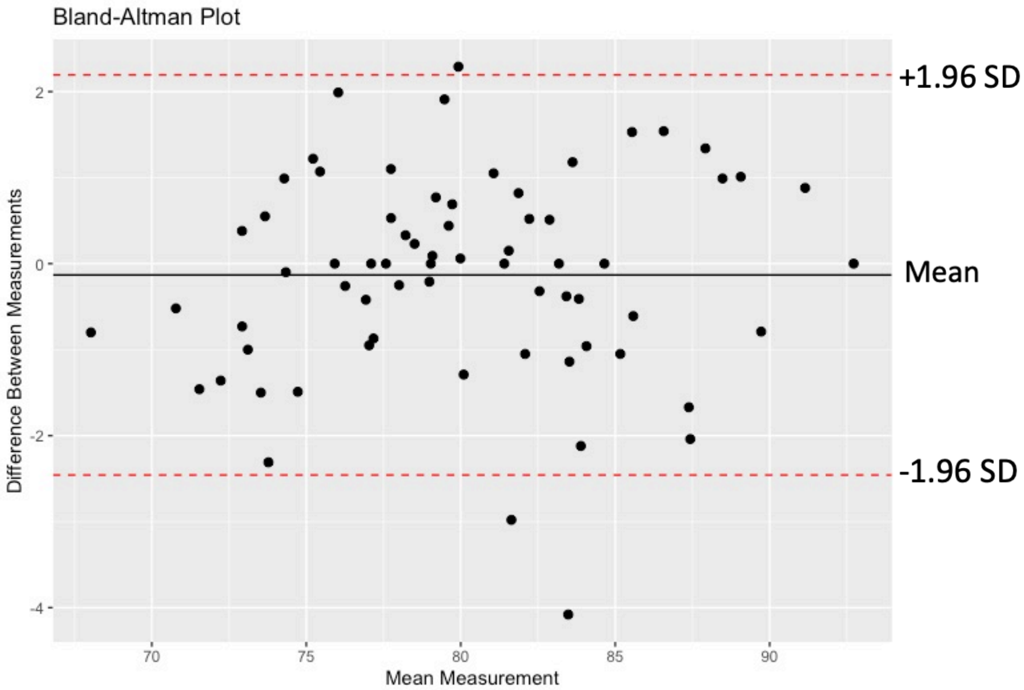
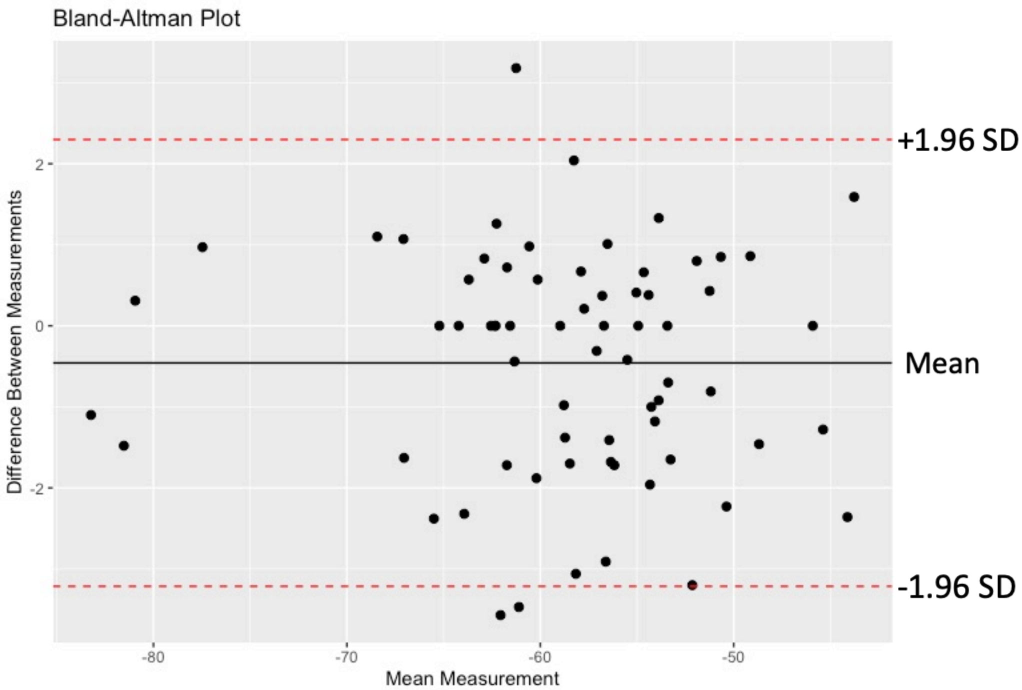


Figure 8.10 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for lower incisor apex (LIA)



Figures 8.9 and 8.10 show that the horizontal and vertical discrepancies were similar for the identification of lower incisor apex by hand tracing and automated tracing.

Figure 8.11 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for porion (Po)

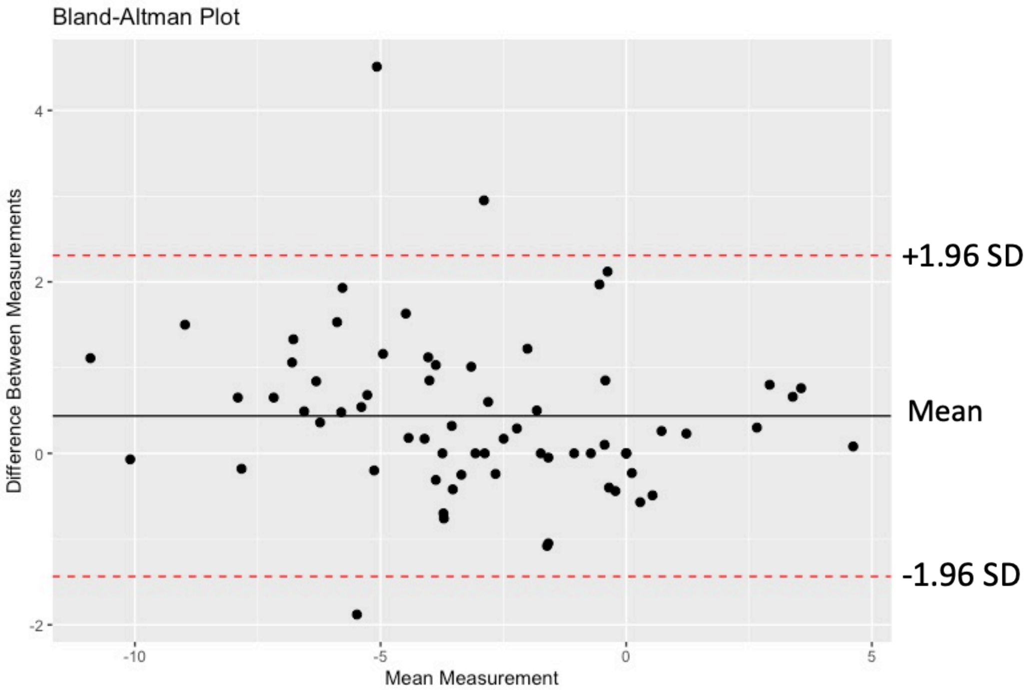
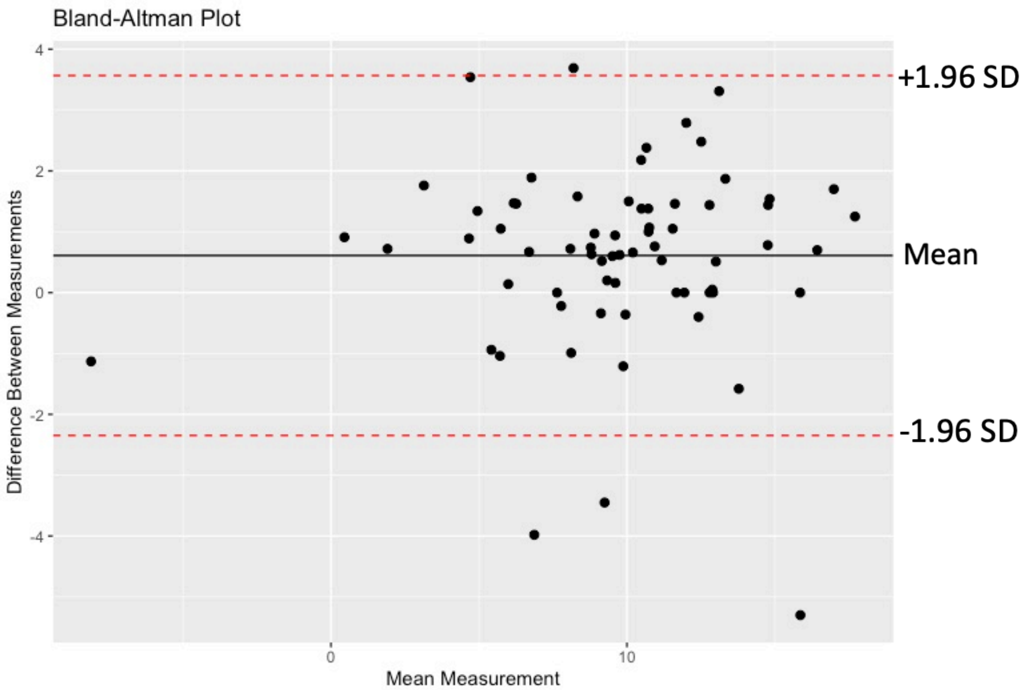


Figure 8.12 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for porion (Po)



Figures 8.11 and 8.12 show that the vertical discrepancy was greater than the horizontal discrepancy for the identification of porion by hand tracing and automated tracing.

Figure 8.13 Bland-Altman Plot showing the inter-rater agreement for x-axis measurements for orbitale (Or)

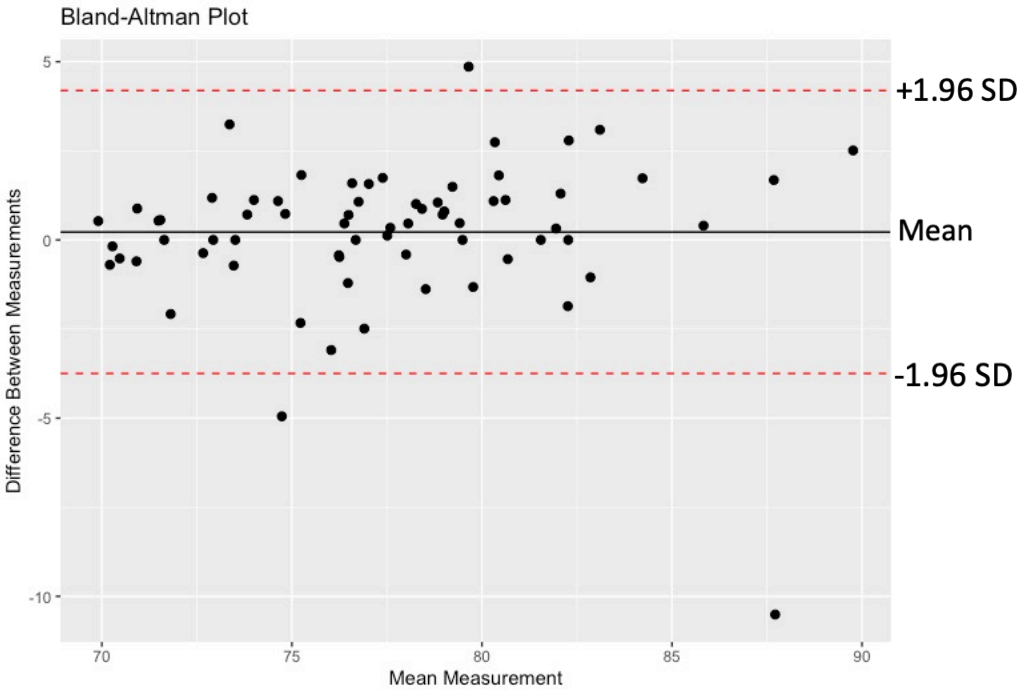
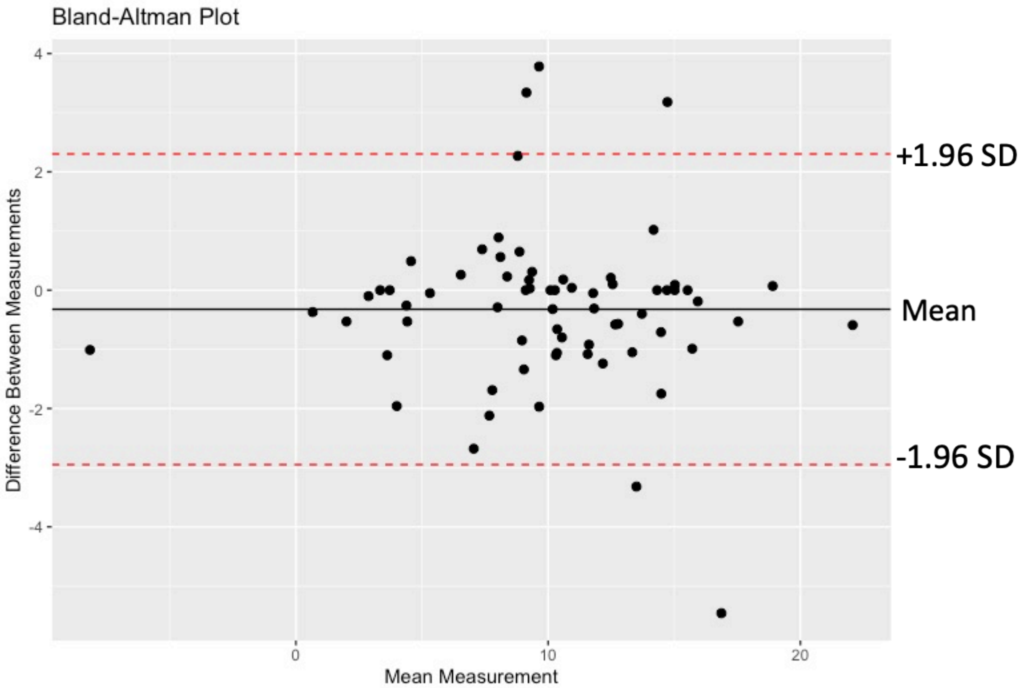


Figure 8.14 Bland-Altman Plot showing the inter-rater agreement for y-axis measurements for orbitale (Or)



Figures 8.13 and 8.14 show that the horizontal discrepancy was greater than the vertical discrepancy for the identification of orbitale by hand tracing and automated tracing.